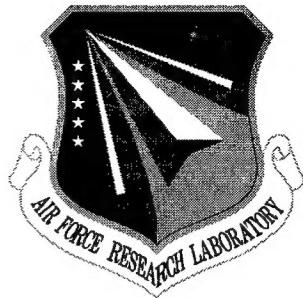


AFRL-SN-RS-TR-2001-37
Final Technical Report
March 2001



WAVE-COUPLED MILLIMETER-WAVE ELECTRO-OPTIC TECHNIQUES

California Institute of Technology

William B. Bridges, Lee J. Burrows, and Uri V. Cummings

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

20010507 059

**AIR FORCE RESEARCH LABORATORY
SENSORS DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK**

This report has been reviewed by the Air Force Research Laboratory, Information Directorate, Public Affairs Office (IFOIPA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

AFRL-SN-RS-TR-2001-37 has been reviewed and is approved for publication.

APPROVED:



NORMAN P. BERNSTEIN
Project Engineer

FOR THE DIRECTOR:



ROBERT G. POLCE, Chief
Rome Operations Office
Sensors Directorate

If your address has changed or if you wish to be removed from the Air Force Research Laboratory Rome Research Site mailing list, or if the addressee is no longer employed by your organization, please notify AFRL/SNDP, 25 Electronic Pky, Rome, NY 13441-4515. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document require that it be returned.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188
<p>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.</p>			
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED	
	MARCH 2001	Final Report Feb 96 - Jun 00	
4. TITLE AND SUBTITLE		5. FUNDING NUMBERS	
WAVE-COUPLED MILLIMETER-WAVE ELECTRO-OPTIC TECHNIQUES		C - F30602-96-C-0020 PE - 63726F PR - 2863 TA - 92 WU - 81	
6. AUTHOR(S) William B. Bridges, Lee J. Burrows, and Uri V. Cummings		8. PERFORMING ORGANIZATION REPORT NUMBER N/A	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) California Institute of Technology Office of Sponsored Programs 1201 East California Blvd, MC 213-6 Pasadena CA 91125		10. SPONSORING/MONITORING AGENCY REPORT NUMBER AFRL-SN-RS-TR-2001-37	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Research Laboratory/SNDP 25 Electronic Pky Rome NY 13441-4515			
11. SUPPLEMENTARY NOTES Air Force Research Laboratory Project Engineer: Norman P. Bernstein/SNDP/(315) 330-3147			
12a. DISTRIBUTION AVAILABILITY STATEMENT APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.		12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) This report details results on two antenna-coupled millimeter-wave electro-optic modulators, the slot-vee antenna-coupled modulator and a 94 GHz directional coupler modulator. The slot-vee-coupled modulator gave 0.17 m-squared per Watt sensitivity, but known defects in the chip imply much better performance is possible. The 94 GHz directional coupler improvements anticipated were not realized due to fabrication problems. A detailed theoretical study of the effects of velocity mismatch on linearized electro-optic modulators was made and the results published. A key result was that directional-coupler based modulators were found to be significantly more easily degraded than some Mach-Zehnder-based modulators. An experimental study of bias drift in electro-optic modulators was made and protons were determined to be the cause. Several inventions were made to reduce or eliminate proton-caused bias drift.			
14. SUBJECT TERMS Electro-Optical Modulators, Millimeter Waves, Linearized Modulators, Optics, Photonics		15. NUMBER OF PAGES 136	
		16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL

TABLE OF CONTENTS

1.0	Executive Summary.....	1
2.0	Introduction and Outline	2
2.1	Introduction	2
2.2	Outline	3
3.0	Directional Coupler Modulators	5
3.1	Background	5
3.2	94 GHz Directional Coupler Modulator Fabrication	8
3.3	RF Waveguide and Coupling Measurements	13
3.4	94 GHz Measurements	19
3.5	Conclusions and Suggestions for Further Work	23
3.6	References.....	24
4.0	Effect of Velocity Mismatch on Modulator Linearization.....	25
4.1	Background	25
4.2	Summary of Appendix A: J.O.L.T. Paper "Bandwidth of Linearized Electro-optic Modulators"	26
4.3	Summary of Appendix B: U.S. Patent 5,886,807.....	30
4.4	Odd-length Directional Coupler Modulators.....	32
4.5	The Y-fed Directional Coupler Modulator.....	36
4.6	References.....	43
5.0	The Slot-vee Antenna-coupled Modulator	45
5.1	Background	45
5.2	Slot-vee Work During the Funding Gap Period	51
5.3	Initial Modulator Feed Design	53
5.4	Measurements with the Original HRL Modulator Chip.....	57
5.5	First Modulation Tests.....	58
5.6	Antenna Coupling Experiments	59
5.7	Modified Coupling Components.....	64
5.8	New Modulator Chips.....	66
5.9	Detection Sensitivity Improvements.....	67
5.10	Final Experiments	68
5.11	Discussion of the K-band Modulator Results	70
5.12	Modulators at 44 and 94 GHz	71
5.13	Conclusions.....	72
5.14	References for the Slot-vee Antenna-coupled Modulator.....	72
6.0	Drift in the characteristics of Integrated Optical Waveguide Modulators	74
6.1	Background	74
6.2	Review of Waveguide Fabrication Techniques.....	75
6.2.1	Titanium In-diffused Optical Waveguides	75
6.2.2	Annealed Proton-exchange Waveguides	76
6.3	Caltech Proton Exchange Waveguide Studies	77
6.3.1	Proton Diffusion Model.....	77
6.3.2	Proton Mobility Measurements	79
6.3.3	Deuteron Exchange Waveguides	80
6.3.4	APE Modulator Drift Measurements	83
6.3.5	Conclusions From Proton Exchange Studies	85
6.4	Proton-free Titanium In-diffused Modulators.....	86
6.4.1	Drift and Aging in Standard Titanium In-diffused Modulators	86
6.4.2	Titanium In-diffused Modulator Drift Experiment	86

6.4.3	Discovery of a Proton-free Process	88
6.4.4	Driftless Titanium in-diffused Modulators.....	90
6.4.5	Future Plans for Low Drift Modulators	93
6.5	Conclusions.....	93
6.6	References.....	94
7.0	Personnel.....	95
8.0	Interactions, Papers, and Patents.....	96
8.1	Meetings and Conferences	96
8.2	Publications.....	96
8.3	Patents	96
GLOSSARY	97
Appendix A	A-1
Appendix B	B-1

TABLE OF FIGURES

Fig 3.1	Electrode mask UVC-el-1 for 11 different antenna coupled directional coupler modulators	8
Fig 3.2	Photograph of the modulator mounting structure	10
Fig 3.3	Modulator chip mounted on the mechanical carrier stage	11
Fig 3.4	Blown up view of the modulator chip and mm-wave slab guide.....	12
Fig 3.5	Slab dielectric waveguide coupling scheme for mm-wave modulators	14
Fig 3.6	Photograph of a "back-to-back" coupling structure	15
Fig 3.7	Plot of dielectric constant correction factor as a function of dielectric constant for a 1 mm thick slab waveguide at 94 GHZ.....	16
Fig 3.8	Photograph of the overall modulator measurement setup	19
Fig 4.1	A reflective traveling wave phase modulator.....	30
Fig 4.2	Gain of a link with a simple DCM using the canonical parameters of Refs 4.2, 4.4 and Append A.....	34
Fig 4.3	Gain for a link using simple DCM with parameters given in Refs 4.2, 4.2 and Append A	35
Fig 4.4	Schematic representation of a YFDCM.....	36
Fig 4.5	Transfer function of a YFDCM.....	37
Fig 4.6	Signal and third-order intermodulation products for a YFDCM with parameters the same as modulators studied in Ref 4.2	38
Fig 4.7	Sensitivity of dynamic range to modulator electrical length for a linearized YFDCM using The Parameters of Ref 4.2	39
Fig 4.8	Gain of Y-fed DCM using the parameters of Ref 4.2	41
Fig 4.9	Gain of Y-fed DCM below $\theta = 1.5 \pi$	42
Fig 5.1	Comparison of vee and slot-vee antennas on a dielectric substrate.....	46
Fig 5.2	Schematic drawing of slot-vee antenna elements overlaying one waveguide of a Mach-Zehnder modulator	48
Fig 5.3	Schematic view of the slot-vee modulator.....	49
Fig 5.4	Schematic drawing of the slot-vee modulator and its microwave feed system.....	50
Fig 5.5	Antenna mask for 5 slot-vee modulators.....	51
Fig 5.6	Assembly drawing of modulator support and feed structure	54
Fig 5.7	Photograph of slot-vee modulator test setup with fiber-input coupling and lens output coupling	55

Fig 5.8	Overall view of the scale model transmission setup showing two slot vee antennas	60
Fig 5.9	Close up detail of one slot vee antenna with coax line feed	61
Fig 5.10	Transmission between slot vee antennas on a 6.7 mm Stycast® $\epsilon_r=30$ substrate	62
Fig 5.11	Same as 5.10 but 28.7 mm thick.....	63
Fig 5.12	Transmission 18-26 GHz with old 20 GHz plug set	65
Fig 5.13	Transmission curve with 24 GHz plugs and top plane	66
Fig 5.14	a) Modulation sideband at 23.472 GHz and carrier (off scale) b) Carrier on scale	69
Fig 5.15	a) Modulation sideband at 23.472 GHz (carrier off scale) b) Carrier on scale	69
Fig 6.1	Conductivity of lithium niobate samples as a function of reciprocal temperature	80
Fig 6.2	Bias voltage drift referred to the initial bias voltage of 7.6 V for an APE MZ modulator	84
Fig 6.3	Fractional drift for the same APE modulator as Fig 6.2 but with an initial bias of 52 V.....	85
Fig 6.4	Drift in bias voltage relative to an initial bias of 5.3 V for a titanium in-diffused modulator made by conventional techniques	87
Fig 6.5	Bias drift in the same modulator as Fig 6.4 but with initial bias of 51 V	88
Fig 6.6	Drift rate for the modulator made by our new dry oxygen process compared to the drifts for APE and conventional Ti in-diffused modulators	91
Fig 6.7	Bias drift in the Caltech process modulator of Fig 6.6 on a greatly magnified scale.....	92

List OF Tables

Table 6.1 Comparison of Experimental and Theoretical Diffusion Depths 78

1.0 EXECUTIVE SUMMARY

This report documents the progress made in improving antenna-coupled modulators for mm-wave, analyzing linearized modulators, and studying and countering bias drift in electro-optic modulators.

We demonstrated for the first time a modulator at 24 GHz using the slot-vee antenna coupled arrangement suggested by F. T. Sheehy in a previous Rome Laboratory contract. The actual performance was measured as $0.17 \text{ m}^2/\text{Watt}$, comparable to other antenna coupled modulators we have built. However, known defects in this particular chip suggest that future performance would be much better, comparable to commercial low microwave frequency modulators.

Our attempts to correct the known defects in the 94 GHz directional coupler modulator first demonstrated in the previous contract were unsuccessful, due to difficulties in developing the skills needed to transfer the fabrication technology to Caltech from Hughes Research Laboratories (now HRL Laboratories). We simply ran out of time before we could get a good chip.

An analysis of the effects of velocity mismatch on linearized modulators was made, with one paper and one patent resulting from this work. We discovered that directional-coupler-based linearized modulators are much more easily degraded by velocity mismatch. We analyzed the Y-fed DCM, an important zero-bias modulator candidate.

Finally, we continued our studies of bias drift in lithium niobate modulators. We have confirmed our hypothesis that protons are the cause, whether introduced intentionally (e.g., in annealed proton exchange waveguides) or unintentionally (e.g., in titanium in-diffused modulators). We demonstrated experimentally that proton diffusion (and the resulting bias drift) can be slowed by substituting deuterium for hydrogen in all process steps. We also developed a titanium in-diffusion process that results in a proton-free chip. This work has not yet been published, but five U. S. Patents have been applied for.

2.0 INTRODUCTION AND OUTLINE

2.1 INTRODUCTION

The work reported here was funded from February 1996 to February 1998 by contract F30602-96-C-0020 by Rome Laboratory. Also reported is work done during an un-funded gap from November 1994, when the previous contract, F30602-92-C-0005, ended through January 1996, and also some work completed after the present contract ended in February 1998.

The general concept of antenna-coupling to a segmented transmission line in a traveling-wave electro-optic modulator was demonstrated several years ago. This technique can be used to overcome the frequency limitations imposed by the finite transit time through a velocity-mismatched transmission system, thus allowing the realization of efficient modulators at mm-wave frequencies. In the previous contract, some new configurations for this general approach were demonstrated (the directional coupler modulator at 94 GHz) and designed (the slot-vee modulator). The present program was undertaken to optimize the 94 GHz directional coupler modulator, and to demonstrate the slot-vee modulator at 20, 44 and 94 GHz.

In addition, we wished to explore how this technique would work with linearized modulators, since these are becoming important in analog photonic systems that require very high dynamic range. Such modulators usually involve the differencing of two channels so that the intermodulation distortion is cancelled, while the signal is not. This cancellation process is known to be very sensitive to the modulator parameters, and we wished to see if the antenna-coupling scheme was compatible.

Another problem that we wished to investigate was the drift in modulator characteristics, particularly with d-c bias applied, that all modulator researchers in the field have observed and for which numerous explanations for the physical source have been put forth. We felt we had a new and compelling physical picture of the underlying physics, and we wished to verify it, possibly developing new methods of countering the drift. These ideas originated during the previous Rome Laboratory contract, with our work on proton-exchange optical waveguides. But we did not have time to make the critical measurements. During the "gap" period, we were able to obtain funding from MIT Lincoln Laboratories for this portion of our work, so that the actual studies were carried out under that program instead of the current Rome Laboratory program.

However, since the same personnel and laboratory equipment were used for those studies, we summarize the results here also.

2.2 OUTLINE

Section 3 reports the progress on the re-designed 94 GHz directional coupler modulator. While the basic modulator was demonstrated in 1993, and reported in January 1994, its performance was severely degraded by defects in its design and fabrication that were discovered later. We undertook to correct the design and fabricate a new lithium niobate modulator for this contract. Since the group at the Hughes Research Laboratories that made the original modulator chip was now assigned to other activities, we had to develop most of the fabrication techniques in-house at Caltech. This section reports the redesign process and the fabrication steps for the new modulator chip. It also describes the optimization of the r-f feed slab waveguide system. Unfortunately, the ultimate measurement of modulation with this chip was unsuccessful. We hypothesize that the failure was due to a residual thin metal film that remained after an etching process that should have removed it. Instead, the chip behaved like a thin-film bolometer in detecting audio modulation on the 94 GHz signal source.

Section 4 outlines the investigation of the effects of velocity mismatch on linearized modulators. Appendix A is a copy of the more detailed publication that resulted from this work, and Appendix B is a copy of a U. S. Patent that also resulted from this study. While various linearization techniques can produce up to 20 dB improvement in dynamic range of a photonic link using the linearized modulator, we found that velocity mismatch can seriously limit the operating bandwidth of some schemes. Schemes based on directional couplers are more adversely affected than some Mach-Zehnder based schemes. The first analysis of the affects of velocity mismatch on the promising Y-fed directional coupler modulator is presented.

Section 5 gives the continued development of the slot-vee antenna-coupled modulator. A 24 GHz version of this modulator was demonstrated successfully, with performance comparable to conventional 24 GHz modulators, provided allowances are made for known fabrication and coupling defects.

Section 6 is a brief summary of the studies of drift in electro-optic modulators that were carried out under the MIT Lincoln Laboratories contract. Experimental results with both proton-exchange and titanium in-diffused modulators are described. Fabrication techniques that result in very low bias drift rates were developed and patent coverage applied for.

Section 7 introduces the personnel that have contributed to this project. Section 8 lists the meetings and conferences attended by project personnel, and the papers, publications, and patents related to the project. A glossary of terms follows Section 8. Appendices A and B are associated with Section 3.

3. DIRECTIONAL COUPLER MODULATORS

3.1 BACKGROUND

Directional coupler modulators (DCM's) suffer a roll-off in sensitivity with increasing frequency due to velocity mismatch, just as Mach-Zehnder modulators (MZM's) do. The -3 dB frequency in modulation sensitivity is about 50% higher for a simple DCM than a simple MZM of the same length and velocity mismatch, but a linearized modulator based on a DCM is much more adversely affected by velocity mismatch than a linearized modulator based on a MZM. This comparison is covered in detail in Section 4 and in Appendix A. In this section we describe experiments designed to show that our antenna-coupling scheme to counter the adverse effects of velocity mismatch works as well with DCM's as with MZM's. This demonstration began several years ago when Finbar Sheehy designed a DCM with an array of simple dipole antennas on the surface of z-cut lithium niobate as part of his Ph. D. research. The design is given in Sheehy's thesis [Ref. 3.1] and also in the Final Report of the previous contract [Ref. 3.2].

Sheehy's modulator design was fabricated at Hughes Research Laboratories (now HRL Laboratories, LLC). Sheehy set up the DCM chip in the laboratory and attempted to make it work. While some asymmetrical modulation sidebands could be observed when the modulator was operated at 94 GHz, he could make no sense of the transfer function of the modulator. This was the situation when he finished his work in 1993.

The problem of understanding what was happening with this modulator was given to a then-undergraduate student, Uri Cummings, as a Summer Undergraduate Research Fellowship (SURF) summer project. The SURF project was successful, and was continued by Cummings as a senior thesis project. Under the SURF project, Cummings was able to deduce that the transfer function of the modulator was correct for a directional coupler, provided one assumed that the directional coupler did not have identical waveguides. That is, the measured transfer function of the dc bias electrodes could be understood if there were a built-in difference, $\Delta\beta$, between the propagation constants of the two guides. (The transfer function of the antenna-coupled sections could not be measured, since no bias contact pads were attached to them in this design). Sheehy's design was intended to have identical optical waveguides in the coupler, so there should have been no built-in $\Delta\beta$.

Understanding the transfer function allowed Cummings to correctly bias the modulator and obtain modulation at 94 GHz. This he did as a senior thesis project, and the resulting performance of $0.2 \text{ [m}^2/\text{Watt]}$ was reported at the PSAA-IV conference in 1994 [Ref. 3.3]. This modulation sensitivity corresponds to a V_{π} of about 40 Volts for a conventional 50 Ohm modulator, which is not all that good for a conventional low-frequency, 9mm active length modulator, but it is good enough to be interesting for a 94 GHz modulator (not commercially available in 1994, nor now). Unfortunately, this particular modulator chip also had a very bad optical transmission, about -20 dB optical loss.

We did not know the origin of the built-in $\Delta\beta$ or the high optical loss at the time the modulator was successfully demonstrated. However, after the PSAA meeting, while we were photo-copying the waveguide masks used by Hughes to fabricate the modulator, we discovered "mouse-bite" defects on the mask: Chips of the mask metal had broken off along the edges of the waveguide regions. These chips resembled a "bite", as if a microscopic mouse had bit into the metal coating, hence the name. Such defects are known in the mask-making industry, and are attributed to various causes. Two such causes are (1) a dirty glass substrate causing adhesion problems when the metal is deposited, and (2) electrostatic discharge when the mask is used over a lithium niobate (pyroelectric) wafer. It turns out that Hughes was experiencing "mouse bite" defects on a number of masks just at the time we discovered them on the particular waveguide mask used in our DCM. There was an on-going argument between Hughes and the mask maker, Align-Rite, with Hughes citing cause (1) when the mask was made at Align-Rite, and Align-Rite asserting cause (2) when the mask was used at Hughes. This argument was never settled, but for us, the damage was already done. Upon close inspection of the modulator under a microscope, we noted that the "mouse-bite" defects in the mask had been faithfully transferred to Sheehy's modulator waveguides, resulting in a significant "bulge" in the waveguide width at every defect. Photographs of these defects were included in the previous contract's Final Report [3.2]. This was the explanation for both the deduced $\Delta\beta$ and the high optical loss: The two waveguides clearly had different defects and thus different β 's, and each defect scattered light out of the waveguide into the surrounding substrate (excess loss).

A new directional coupler waveguide mask, identical to the original one designed by Sheehy was fabricated on z-cut lithium niobate (but without the “mouse bites”), and received in mid 1994. New modulator chips were fabricated with this mask by Hughes and delivered to Caltech. These chips exhibited reasonable optical loss (-8 dB), and the transfer functions of the bias sections were now symmetrical about zero bias (as a DCM with $\Delta\beta = 0$ should be). However, the transfer functions exhibited very little change in transmission with applied voltage. After some investigation, it turned out that the original directional coupler design parameters determined by Dr. Cel Gaeta at Hughes Research Labs (unpublished) were incorrect. Following Gaeta’s data, Sheehy had used the wrong dimensions in his directional coupler modulator design. The waveguides in the coupler were too far apart, resulting in too little coupling. The bias sections of the coupler could not adjust the modulator to the proper condition for the antenna-coupled sections to work. In fact, the only reason that Cummings was able to obtain the results reported in Ref. 3.3 was due to the $\Delta\beta$ and light scattering introduced by the mouse bites! The new modulator, without the mouse bites, was too insensitive to bias properly.

Fortunately, this error had also been discovered at Hughes, and Gaeta was already in the process of making new measurements of the directional coupler properties. This was the situation at the end of the previous Rome Laboratories contract in November 1994, and was reported in detail in the final report [Ref. 3.2].

We continued to work on this modulator during the funding “gap” between that contract and the present contract, albeit more slowly than before or since. Using Gaeta’s new directional coupler parameters, Cummings designed new waveguide and electrode masks. The new electrode mask layout (Denoted UVC-el-1) is shown in Fig. 3.1 (repeated from Ref. 3.2). Eleven different antenna-coupled DCM’s are on the mask, along with some simple directional coupler test sections. The modulators have various combinations of antenna sections and dc bias sections (heavy lines). The gap between the bias electrodes is not resolved on this figure). The masks were fabricated (by Align-Rite, but without the mouse bites) and sent to Hughes Research Laboratories for modulator fabrication.

The antenna-plus-transmission-line-segments are arrayed in the center of each modulator. The large black rectangles at the ends are the bias sections, which are connected to contact pads at

the right and left edges of the mask. The gap between the two halves of the bias electrodes is so small (a few microns) that it is not discernable on this figure. Various combinations of bias electrode length and position, and antenna array length and position were included. The shorter horizontal traces at the top and bottom of the pattern are simple transmission lines for conventional DCM's.

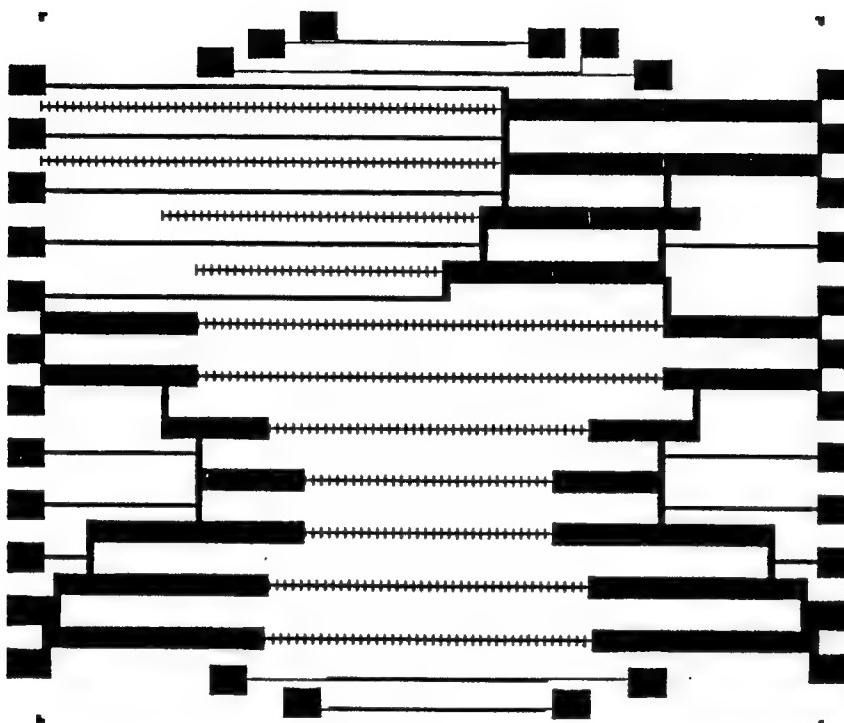


Fig. 3.1 Electrode mask UVC-el-1 for eleven different antenna coupled directional coupler modulators. [From Ref. 3.2]

3.2 94 GHZ DIRECTIONAL COUPLER MODULATOR FABRICATION

The present contract began in mid-February, 1996. We were suffering a bit of dislocation at this time, having just moved all our offices and laboratories from the Watson building to the new Gordon and Betty Moore Laboratories building at Caltech. HRL had fabricated the directional coupler titanium in-diffused waveguides and applied a silicon dioxide buffer layer. The wafer was returned to Caltech in May to have the electrodes applied. Since this was our first experience with making the electrodes ourselves, we proceeded with some caution and

preliminary experiments. We used Prof. Axel Scherer's laboratory facilities for this processing. Prof. Scherer routinely uses negative photoresist while HRL uses positive resist and a "lift off" gold process (which is why we had been copying and reversing the various masks when we discovered the "mouse bites"). We ran tests in spinning-on and evaluating negative resist on glass slides to make sure we understood its characteristics. We found that we could obtain good results, but that the exposure time and level were much more critical than with the positive resist we had previously used (at Caltech and HRL).

Using negative resist, a uniform gold layer is evaporated over the entire surface of the lithium niobate wafer. Then the photoresist is spun on, dried, and exposed with the electrode mask. The photoresist is developed and washed away everywhere except over the gold where the modulator electrode pattern is desired. Then the rest of the gold is etched away in the exposed regions with a reactive ion etching (RIE) process.

Our use of negative photoresist created some difficulty in the fabrication process that may be relevant to the ultimate results obtained for this modulator. With the negative resist process we used, the edges of the metallic regions were less well defined than when positive resist is used. This became a concern in the region between the two electrodes forming the traveling wave transmission line and the region between the halves of the bias electrodes. These gaps are only a few microns. After the first etching step, we measured some conductivity between these electrodes, indicating there was still some gold bridging the gap. This required spinning on more resist, exposing the mask again (with the difficulties of re-aligning it to exactly the same position, developing, and continuing the etching process. The consequence of over etching would be removal of the traveling wave electrodes themselves, so it was important to get the exact total etching time correct. The time was judged to be correct if there were high dc resistance between the two electrodes. However, the resulting structure likely had continuously-varying thickness "filets" across the gap. This may have greatly upset the distribution of transverse electric field between the electrodes.

Since we had used a simple evaporated gold coating without a chromium underlayer (in order to allow the gold to be more easily etched off), the result was a very fragile gold electrode pattern. To protect this fragile layer, a silicon dioxide over-coating was applied. The modulator

substrates were returned to HRL for this coating, since we were not yet experienced in using the silicon dioxide sputtering apparatus in Prof. Scherer's lab. When the wafer was received back from HRL, another photoresist layer was spun on and exposed with a mask that gave electrical access to the bias pads in the underlying gold electrode layer. Using this mask, we etched through the silicon dioxide protective layer. The modulators on the wafer were checked for dc short circuits. None were found. The wafer was returned to HRL to be sawed into two 1" by 2" substrates, which were then polish-cut on the ends.

A mechanical support structure for the modulator substrate was designed and fabricated by the Caltech central machine shop [Fig. 3.2]. The substrate is mounted on a carrier oriented in a

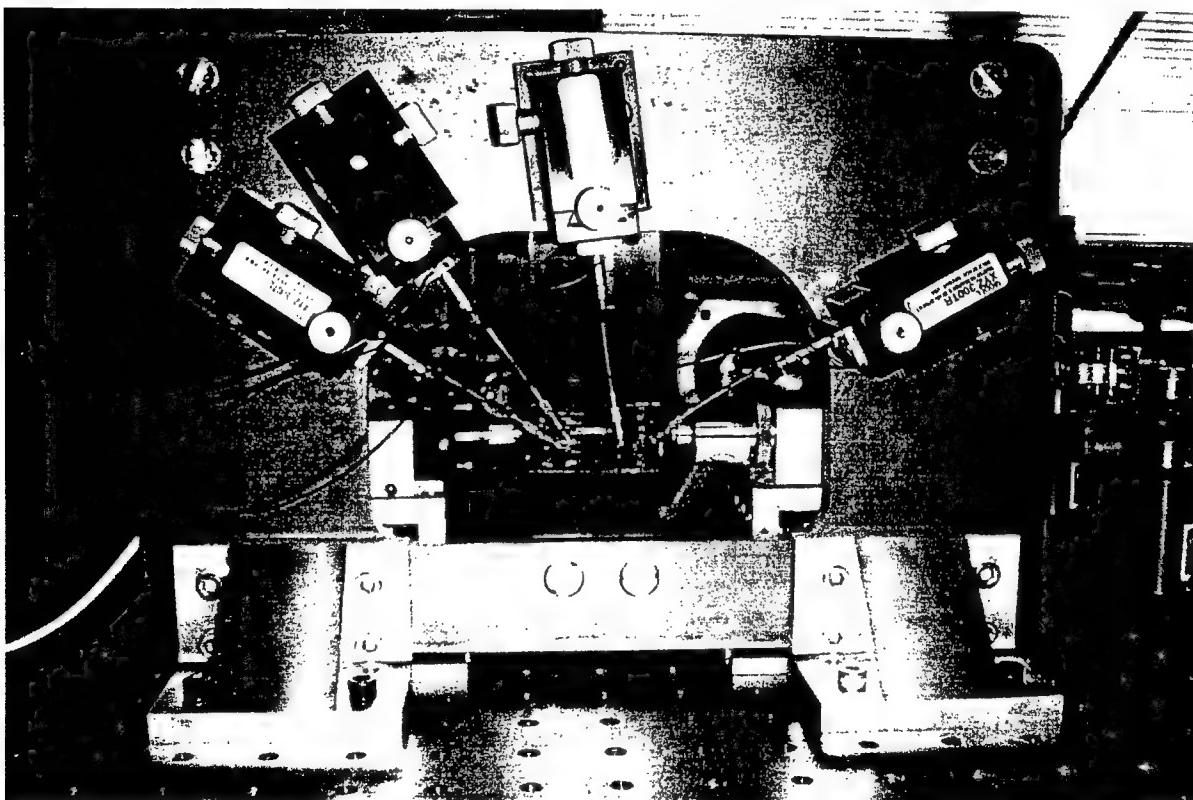


Fig. 3.2 Photograph of the modulator mounting structure.

vertical plane. Both sides of the substrate are accessible so that the dc bias voltages may be applied with probes on the front side and the mm-wave illumination applied with a slab

waveguide from the back side. A vertical steel plate surrounds the carrier so that the magnetic-based probe holders may be positioned on it. A closed-up of the modulator mounted on its carrier is shown in Fig. 3.3.

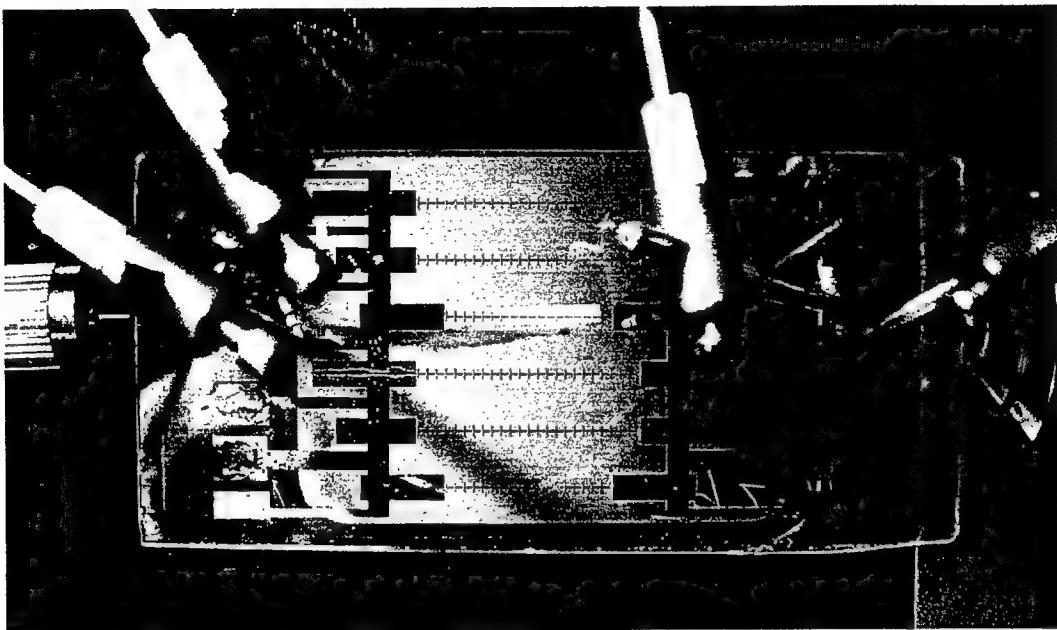


Fig. 3.3 Modulator chip mounted on the mechanical carrier stage.

A further enlargement is shown in Fig. 3.4. In this figure, the dipole-plus-transmission-line-segments are seen more clearly against the end of the lithium niobate slab waveguide contacting the modulator chip from behind (white rectangle). The large rectangles at either end of the dipole array are the bias electrodes. (The gap in between the bias electrodes and the transmission line electrodes is too small to be seen, even in this blow-up).

Optical transmission measurements were initially made at 633 nm, and the transmission appeared to be good (although multimode, since the waveguides were designed for 1.3 micrometer wavelength). We were awaiting the delivery of some GRIN-lens-terminated fibers to do the optical input and output coupling at 1.3 micrometers.

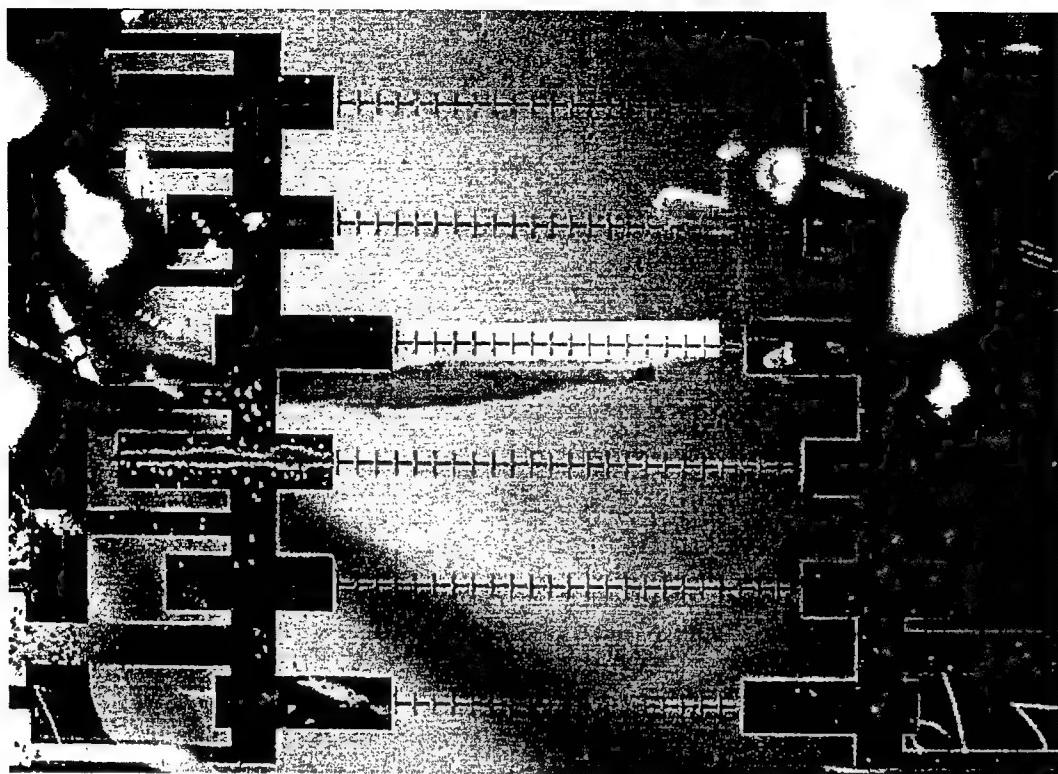


Fig. 3.4 Blown up view of the modulator chip and mm-wave slab guide.

At this point in the investigation, a most unfortunate accident occurred to Uri Cummings, the student carrying out the DCM studies. He broke some small bones in his wrist in a skiing accident, and had to have his hand immobilized in a cast. These bones proved very slow to heal, and the cast remained on for several months, followed by physical therapy to regain the use of his hand. Very little work could be accomplished in the laboratory, so Mr. Cummings worked on the computer calculations for the linearized modulator studies (described in Section 4 of this report) and the manuscript for the publication of these studies (see Appendix A). (He could still type one-handed).

While Cummings was recuperating, we were working on improving the entire laboratory setup by converting to single-mode, polarization-preserving fiber with FP/APC connectors throughout the lab, so that every source, detector and input/output fiber could be interchanged easily with minimum realignment. This work was done by graduate student Lee Burrows. While most of the changeover went smoothly, the input/output coupling fibers proved very troublesome. When

we finally received the GRIN-lens-terminated fibers, measurements proved them to be inadequate in their optical quality, and were mechanically unusable. We also experienced a great deal of difficulty using commercial fiber chucks from Newport Corporation (non-axial position and drift in alignment), and eventually we abandoned them in favor of a Caltech-designed and fabricated fiber holder. Eventually, we settled on butt-coupled fibers in the Caltech chucks as the input source for all our experiments, for both the DCM of this section and the Slot-vee experiments described in Section 5 of this report.

Having regained the use of his hand, Cummings set up the modulator and made transmission measurements at 1.3 micrometers with the new fiber-coupled components. The transmission loss averaged about -8 dB for the several modulators on the substrate. While not "commercial quality," it is much better than the -20 dB of the chip with the "mouse bite" waveguides. The dc electrodes sections were biased and exhibited the proper transfer functional form, indicating that Gaeta's new values for the coupling parameters were correct and had been used correctly in the design of the modulator. Bias was applied for several days to assure that drift would not be a problem with these chips. The bias point appears to be sufficiently stable for our measurements.

3.3 R-F WAVEGUIDE AND COUPLING MEASUREMENTS

Our 94 GHz Varian VRC-2113B23 klystron was tested, using a power supply borrowed from Prof. David Rutledge's group. (Our Harris klystron power supply appears to have quit permanently, perhaps a blessing in disguise, since its earlier "protection circuit" failure had burned out our only 60 GHz klystron!) The 94 GHz tube still measures approximately 80 mW output over its 90.75 to 97.75 GHz tuning range.

The mm-wave coupling structure used to excite the antennas on the modulator utilizes a tapered, low dielectric constant dielectric slab waveguide "expansion fan" to increase the width of the slab waveguide from the 0.100 inch width of WR-10 waveguide out to the length of the antenna array on the modulator, approximately 10 mm. The wide end of the fan is butt-coupled to a lithium niobate slab waveguide ($\epsilon_r = 28$) through a "quarter-wave" section of intermediate dielectric constant to minimize reflections. The lithium niobate waveguide is cut in a wedge of the correct angle, and is glued to the backside of the modulator chip. This coupling scheme is

shown schematically in Fig. 3.5. It was used successfully by Sheehy [Refs. 3.1, 3.2] and by Cummings [Ref. 3.3], but not truly optimized.

To optimize the coupling of the mm-wave modulation into the modulator, reflections at the junctions between the various slab waveguide sections must be minimized. This must be done by trial and error, since no accurate theory exists for the metal-waveguide-to-slab wedge

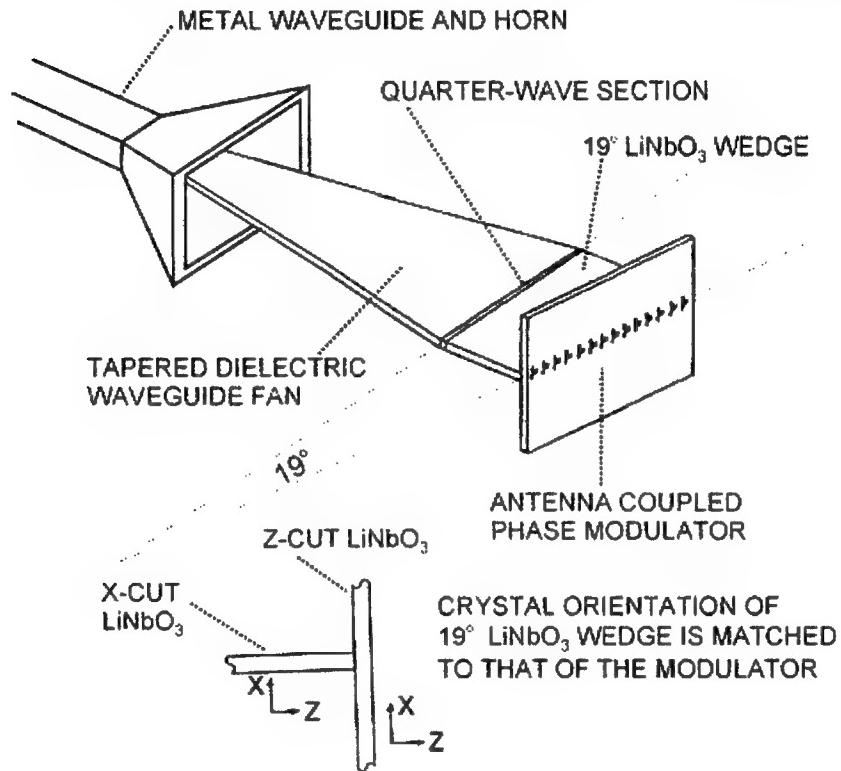


Fig 3.5 Slab dielectric waveguide coupling scheme for mm-wave modulators

junction, nor for the slab-to-slab junctions on the “quarter-wave” matching section. To carry out this optimization, we made transmission measurements of the two such structures, coupled “back-to-back.” We can then assume that the transmission loss from WR-10 waveguide to the modulator would be approximately half the loss (in dB) of the overall transmission. A photograph of a back-to-back slab waveguide arrangement is shown in Fig. 3.6. Two polypropylene tapered slab waveguides are coupled to a uniform section of lithium niobate through “quarter-wave” thicknesses of an intermediate dielectric constant material, Stycast®.

In the measurement setup, the fanned-out sections of slab waveguide were inserted into W-band horns. The overall transmission was measured and compared to the metal waveguide

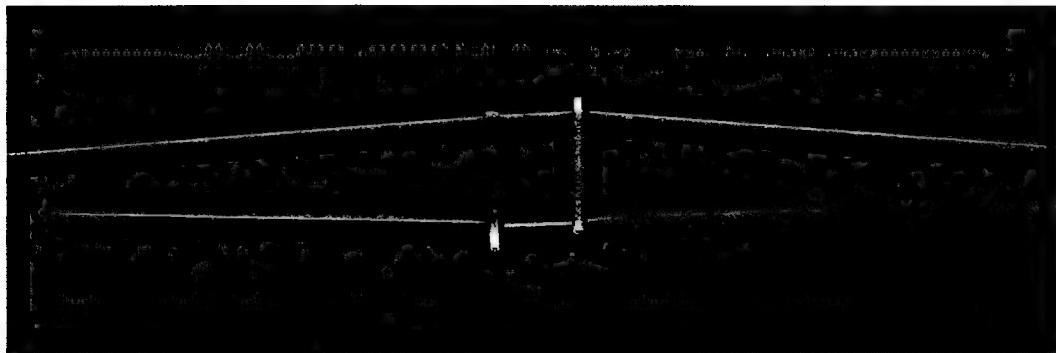


Fig. 3.6 Photograph of a "back-to-back" coupling structure.

flange-to-flange signal. To remove the effects of detector non-linearity, a precision attenuator was used to keep the detector signal level the same for both the transmission through the entire structure and the transmission through just flange-to-flange WR-10.

Many combinations of fan dimensions, and matching section dielectric constants and thicknesses were evaluated. Polypropylene ($\epsilon_r = 2.2$) 1mm thick was originally used for the expansion fans. The 1mm thickness is a good match for the 1.25 mm height of WR-10 waveguide and gives a reasonably well-confined mode on the slab. In the case of infinite plane-wave transmission through a dielectric interface, a "quarter-wave" matching layer should have a dielectric constant equal to the geometric mean of polypropylene and lithium niobate, $(2.2 \times 28)^{1/2} = 7.8$, and should have a thickness of $0.25 \times (3.2 \text{ mm}) / \sqrt{7.8} = 0.28 \text{ mm}$ (or an odd multiple thereof). Of course, here we are dealing with slab waveguides where the fields are not fully confined to the dielectric material. The effective dielectric constant of the guide will be reduced in all cases. Simple slab waveguide theory [Ref. 3.4, for example] allows us to calculate the change in effective dielectric constant. For the polypropylene guide 1 mm thick at 94 GHz, the effective dielectric constant is $0.60 \times 2.2 = 1.32$. For the lithium niobate guide 1 mm thick at 94 GHz, it is $0.916 \times 28 = 25.6$. For the intermediate dielectric constant material in the vicinity of 8

the correction factor is about 0.72. These factors are for TM_0 propagation in the slab, as we have here. For TE_0 propagation, the correction values are somewhat larger. Figure 3.7 shows the dielectric constant correction factor for both TE_0 and TM_0 modes as a function of the slab dielectric constant for 1mm thick slab waveguides at 94 GHz.

One might think that, knowing these correction factors for the effective dielectric constant, that the “quarter-wave” matching formula could be recalculated for an ideal match. Unfortunately, in addition to the change in propagation velocity and wave impedance, there will be a mismatch in

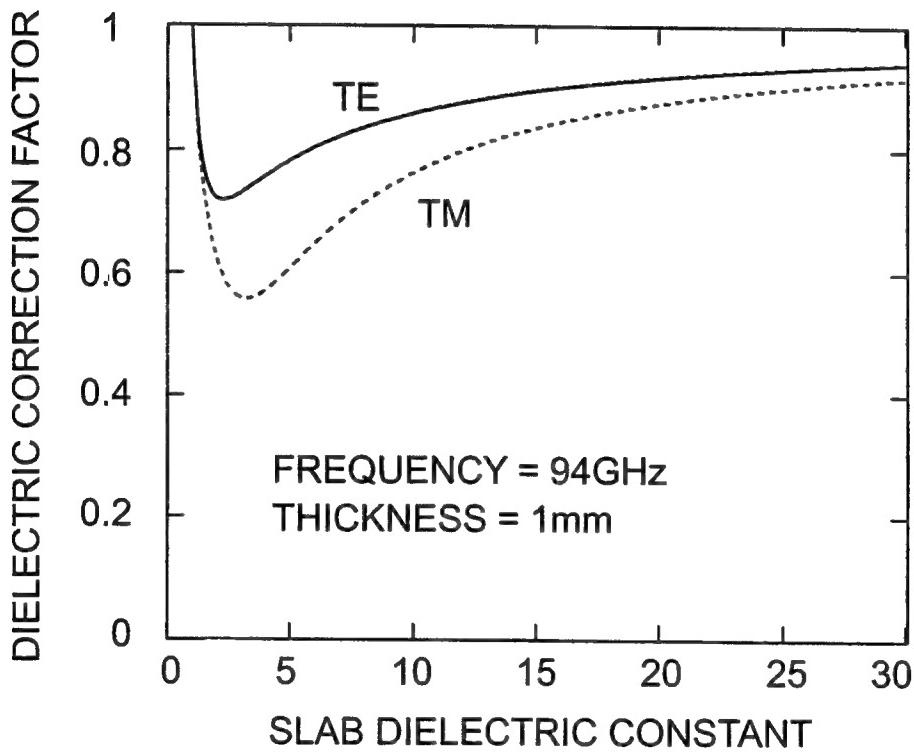


Figure 3.7 Plot of dielectric constant correction factor as a function of dielectric constant for a 1 mm thick slab waveguide at 94 GHz.

the spatial distribution of the modes in the three regions, so that there will be a “discontinuity” reflection, even if the mode velocities and wave impedances are correctly arranged for “quarter-wave” matching. For these reasons, we proceeded largely by trial and error, using the materials at hand rather than ordering special values of dielectric constant.

An additional problem arose as the back-to-back measurements proceeded, that of mechanical instability. Polypropylene is a flexible plastic material, and it proved difficult to maintain a flat, rigid structure while making the measurements. Sheehy had also had this problem, but decided to simply "live with it." This time we decided to change materials, and chose fused silica (1mm thick) for the expansion fans. Fused silica has a dielectric constant of $\epsilon_r = 4.0$ at mm-wave frequencies, so that the optimum (infinite plane wave approximation) matching dielectric constant would be $(4 \times 28)^{1/2} = 10.6$, and a quarter wave would be $0.25 \times (3.2\text{mm}/\sqrt{10.6}) = 0.25$ mm thick. The correction factor for a 1 mm thick slab at 94 GHz is 0.572, so the effective dielectric constant is 2.3. For a matching slab 1 mm thick at 94 GHz with dielectric constant about 10, the correction factor is 0.765 for the TM_0 mode.

It is difficult to make actual quarter-wave layers because they are so thin. The dielectric material we used was Emerson and Cuming Styrocast® artificial dielectric. This material is made from a high dielectric constant powder suspended in a low loss plastic matrix. The dielectric constant is controlled by the fraction of the powder used. It is not a strong material, and thin sections break easily. The most successful method we found to make thin layers was to cut as thin a piece as possible with a saw, then glue it to the dielectric fan and lap it to the desired thickness with sandpaper on glass. Even so, the material would break off if we attempted to make sections actually one quarter wavelength in thickness. Typically, we used odd numbers of quarter wavelengths to gain some material strength. Unfortunately, this increases the accuracy with which the material thickness must be fabricated, whether trying to make a thickness specified from theory, or simply varying the thickness in the trial-and-error mode.

The best results obtained for matching from WR-10 to a polypropylene dielectric fan, through a matching layer to a lithium niobate section and back again were flange-to-flange losses of 11.6 to 12.2 dB. We note that an equivalent length (about 40 cm) of WR-10 waveguide would have a loss of about 1.5 dB. Two configurations gave this result. Two different matching layer configurations gave this same range of losses:

- (1) 1.5 mm of $\epsilon_r = 5$ material
- (2) 1.0 mm of $\epsilon_r = 10$ material.

These values do not agree with our simple picture of “quarter-wave layers.” For 94 GHz the free space wavelength is 3.19 mm. For case (1), correcting for $\epsilon_r = 5$ gives a guide wavelength of $3.19/\sqrt{5} = 1.43$ mm, and the 1.5 mm layer becomes 1.05λ thick, not an odd multiple of $\lambda/4$. Further correcting for the effective dielectric constant in a slab waveguide of 1mm thickness from Fig. 3.7, gives an effective dielectric constant of $\epsilon_r = 0.61 \times 5 = 3.94$, and a guide wavelength of 1.61 mm, and an electrical thickness of 0.95λ , still not an odd multiple of $\lambda/4$. Case (2) is equally puzzling. For $\epsilon_r = 10$, we have a “bulk” guide wavelength of $3.19/\sqrt{10} = 1.01$ mm, so the 1.0 mm layer is almost exactly 1λ thick. Figure 3.7 gives a correction factor of 0.76 so the effective dielectric constant is actually 7.6 and the guide wavelength is $3.19/\sqrt{7.6} = 1.16$ mm for an electrical length of 0.86λ , again, not an odd multiple of $\lambda/4$. We should not expect exact values of $\lambda/4$, however, since there are discontinuity effects at both ends of the matching layer that will produce reflections in addition to those at the surfaces themselves. The most we can say is that the values obtained in cases (1) and (2) above gave the minimum reflection of all those we tried.

We also tried fused silica as the material for the fans instead of polypropylene, primarily because of the higher rigidity of the fused silica. However, the transmission losses were a few dB higher, so we continued to use polypropylene.

An attempt was made to reduce losses that originate as radiation from the dielectric discontinuities by layering copper foil on each side of the slab waveguide assembly. This idea reduced loss in the case of the microwave feed for the slot-vee antenna (see Sec. 5), but it did not do so here. Transmission losses increased about 3 dB with the copper foil in place. Actually, this is the direction we would expect from theory, since the currents induced in the copper walls add loss compared to the negligible loss in the exterior dielectric (air) when they are absent. We can say that the losses in the two copper walls are 3 dB worse than the radiation losses from the discontinuities in the slab guide. The foil-clad waveguide has the merit that the modes can be calculated exactly, and theoretical values of dielectric constant and length can be obtained for the proper match. And the transverse distributions of the TM_0 mode are forced to match by the confining metal conductors. However, the additional loss observed experimentally seemed to argue against using the foil clad waveguide, so we abandoned this approach.

In any event, we estimate that the combination of fan and matching layers given in case (1) or (2) above would exhibit the order of 6 dB loss between the WR-10 waveguide and the bottom surface of the modulator.

3.4 94 GHZ MEASUREMENTS

The overall setup for the actual modulation measurements is shown in Fig. 3.8. The 94 GHz klystron is on the left, mounted in a large finned block. A cooling fan blows on the fins. A ferrite isolator is used to eliminate reflections back into the klystron. Following the isolator, a -20 dB cross-guide directional coupler is used to sample the klystron output for the precision absorption wavemeter. The main signal continues through WR-10 waveguide to a mm-wave horn and the dielectric slab fan (obscured behind the modulator support structure in this photograph).

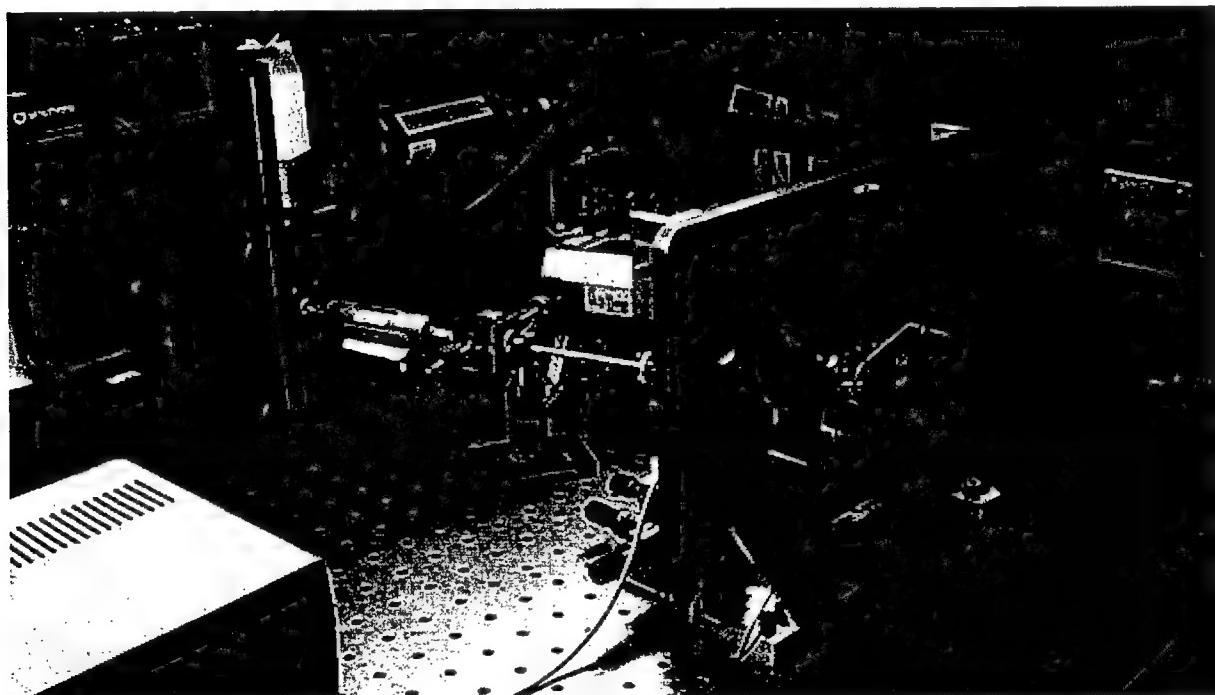


Figure 3.8. Photograph of the overall modulator measurement setup.

The optical power to the modulator at 1.3 microns was input via a butt-coupled fiber mounted in a precision 3-axis motion. The modulator output was imaged with a microscope objective lens

onto a scanning Fabry-Perot (SFP) interferometer (right edge of Fig. 3.8) used to detect modulation side bands. (Recall that we do not have a photodetector with 94 GHz response. Instead, we evaluate the modulator by looking for sidebands in the optical spectrum). The free spectral range of the SFP was 2.00 GHz, so there would be the order of 50 times aliasing in observing the sidebands. However, the frequency of the klystron was measured with an Alpha Industries model W551 wavemeter, with absolute accuracy of 0.2%, (about 200 MHz at 94 GHz) so we are able to tell fairly well where the klystron-induced sidebands should occur within the 2.00 GHz display range. The measurement technique was essentially the same as the measurement that had proven successful before [Ref. 3.3], except now the mechanical structure was much more stable, and the new modulator chip exhibited lower optical loss and a reasonable transfer function compared to the “mouse bite” modulator of Ref. 3.3.

Unfortunately, the expected modulation sidebands were not observed. Many variations on the measurement were made (described below), but none produced sidebands visible on the oscilloscope trace of the SFP output. In the course of these measurements, a curious phenomenon was observed, one for which we have no explanation, only speculation, and that is also described below.

One difference from the old measurement was the source laser. We began by using an ATX model 1.3AMPU 1.3 micron Nd-YAG laser, one capable of about 150 mW output. The old experiment used an Amoco Laser Company (later, ATX) model ALC 1320-EHS laser, that produced only 5 mW maximum, the only laser we had at the time. In an ordinary photonic link, the link gain is proportional to the laser power, which is why we switched to the higher power laser. However, in our detection scheme, we wish to measure the relative height of the sideband and the carrier, as detected on the SFP display. This ratio is independent of laser power, as long as the modulation sideband is visible above the oscilloscope noise level. However, the sideband must also be somewhat above the “skirt” of the carrier as displayed on the oscilloscope screen. (The “skirt” results from the finite width of the SFP filter; see Fig. 5.13 in Section 5, for a typical output of the ATX 1.3AMPU laser.) To minimize the interference from the skirts of the carrier, it is best to tune the modulation frequency to a value exactly half way between the aliased “carriers” on the SFP. For a 2.00 GHz SFP free spectral range, this means the modulation frequency should be chosen at odd numbers of GHz: 91.00, 93.00, 95.00, or

97.00 GHz for the tuning range of our klystron. Operation at 92.00, 94.00, or 96.00 GHz would result in modulation sidebands that fall exactly on the aliased carriers on the display and would be completely undetectable.

An unfortunate disadvantage of the new high power laser was that it seldom operated with a single optical mode, typically oscillating at three or four optical frequencies simultaneously. (It was supposed to be a truly single-frequency laser, and often was when it was new. But by the time we were making measurements on this modulator, the components within the laser had changed so that it now never operates at a single frequency). As a consequence, the SFP 2.00 GHz display was now cluttered with several aliased "carriers" from all the frequencies output from the laser. (Again, see Fig. 5.13b.) Each of these carriers should display sidebands, but it is now more difficult to find a modulation frequency that would place any sidebands at the low points among all the aliased carrier skirts.

After some frustrating days with the high power laser, we returned to use the 5 mW output of the original laser. This made the oscilloscope display less cluttered, but also failed to produce any visible sidebands. And, of course, we now ran the risk that the sidebands were actually below the oscilloscope noise level.

The next attempt at improving the measurement involved building new higher-gain, lower-noise amplifiers for the photodetector. We had to convince ourselves that we were limited by the photodetector amplifier noise, not the built-in oscilloscope amplifier noise. This proved a bit tricky to do. The eventual configuration used a special op-amp as a trans-admittance amplifier for the photodetector, one with very low equivalent current noise rather than the more usual low equivalent voltage noise. This is done because the photodiode source looks like a current source rather than a voltage source. The op amp we selected was the Analog Devices 795, the lowest current noise op amp we could find. The feedback resistor in the first stage was 20 megOhms, so the output of the amplifier is 20 Volt per microamp input. With this amplifier, we were assured that the noise level viewed on the SFP output on the oscilloscope was amplifier-noise-limited. However, with this improved amplifier, we were still unable to discern modulation sidebands.

In an attempt to further improve the sensitivity of the setup, we fed the SFP amplifier output to a “lock-in amplifier” (audio phase-synchronous detector) as well as the oscilloscope. The klystron power supply had built-in square-wave modulation capability at 1000 Hz, and this was enabled. Thus the DCM was now being fed a W-band signal that is 100% on-off modulated at 1000 Hz. With this arrangement, we hoped to use the very narrow equivalent bandwidth of the synchronous detection process. Of course, we now had to scan the SFP very slowly to accommodate the long integration time of the lock-in. Typically, we scanned the SFP manually, observing the deflection of the meter on the lock-in. Using this technique, we were still unable to find modulation sidebands at locations in the SFP frequency scan where we predicted them to be from the precision W-band frequency meter. From these “heroic” detection improvements on the original measurements of the “mouse-bite” modulator, we are forced to conclude that the present modulator chip is at least three orders of magnitude less sensitive than the old modulator, for reasons we do not understand, but are due to faulty chip fabrication in some way.

It was in the course of making the synchronous detection measurements that we discovered a new and, so far, unexplained phenomenon. We were able to see repeatable synchronously-detected outputs from the SFP detector/amplifier combination when the SFP was tuned to the steep slope of an aliased carrier signal!

The behavior was as if the applied mm-wave modulation were frequency-modulating the optical signal at 1000 Hz, and then we used “slope detection” of that FM by tuning the optical filter of the SFP to have its steep slope tuned to the optical carrier frequency. We checked to see that this phenomenon was, indeed, coupled into the modulator via the mm-wave slab waveguide, rather than by stray pickup in the bench wiring. It was.

Our only hypothesis at this point is that the modulator chip itself is somehow “detecting” the W-band signal, and using the “detected” output (a 1000 Hz square wave) to produce phase modulation (which would be indistinguishable from frequency modulation in this experiment). It may be possible that in the process of evaporating and etching the electrodes several times on this one chip (using the negative resist process in Prof. Scherer’s lab, rather than the positive resist “gold lift-off” method that HRL used on the original modulator) we had inadvertently left a thin conductive layer bridging the transmission line electrodes, which acted as a “thin-film

bolometer" that "shorts out" the mm-wave signal, but heats and cools rapidly enough to phase modulate the underlying optical waveguides. Admittedly, this explanation is grasping at straws. The only way to check for sure is to make another modulator chip by the original process (or by successfully developing the negative resist process) and find successful results. Unfortunately, we did not have time to do this under this program (nor, in fact, in the months following its completion).

3.5 CONCLUSIONS AND SUGGESTIONS FOR FURTHER WORK

The basic idea of making an antenna-coupled mm-wave modulator based on a directional coupler optical structure was proven during the previous contract [Ref. 3.3]. The present program was aimed at improving and optimizing the structure (mainly by removing serious defects that were discovered in the original design). Unfortunately, we had to switch fabrication methods to be able to make the new modulator at Caltech (since HRL had subsequently stopped making modulators in-house). And we can only conclude that we have not been successful in developing these techniques.

In the meantime, a new directional coupler structure has come to light that would be a much better choice for an antenna-coupled structure, the Y-fed DCM [Ref. 3.4]. This structure is described in more detail in Section 4 of this report, since it is also a good candidate for linearization. The Y-fed DCM is simply a DCM with equal optical powers fed to the two input arms, rather than all of the optical input fed into one arm. Due to this symmetrical feed, the transfer function changes dramatically. With no voltage applied, the output is also symmetrical from the two output guides (as you can readily imagine by the simple geometric symmetry of the device and its feed). When voltage is applied to electrodes on the directional coupler section, the output in one arm will increase and that in the other arm will decrease, or vice versa, depending on the polarity of the applied voltage. Thus, "zero" is automatically the correct bias for this modulator. This means that an antenna coupled modulator need only the antenna-plus-transmission-line-segments to couple the r-f into the modulator. No extra bias electrodes are required (or desired, for that matter). This structure is a "natural" for antenna-coupled devices.

We may have the opportunity to test this version of an antenna-coupled DCM in the near future. HRL has a new DARPA program to study very high performance A/D conversion using an

optical front end in the sampling process. One of the subtasks potentially involves a 50 GHz fixed-frequency modulator. Since Caltech is a subcontractor to HRL on this program, and since HRL is now back in the modulator fabrication business, it may be possible to fabricate an antenna-coupled 50 GHz (or a 100 GHz) Y-fed DCM that we can evaluate at Caltech.

3. 6 REFERENCES

- 3.1 F. T. Sheehy, "Antenna-Coupled mm-Wave Electro-Optical Modulators and Linearized Electro-Optic Modulators," Ph. D. dissertation, California Institute of Technology, June 1993.
- 3.2 W. B. Bridges, L. J. Burrows, U. V. Cummings, R. E. Johnson, and F. T. Sheehy, "60 and 94 GHz Wave-Coupled Electro-Optic Modulators," Final Technical Report RL-TR-96-188 on Contract F30602-92-C-0005 with Rome Laboratories, September 1996.
- 3.3 W. B. Bridges, U. V. Cummings, F. T. Sheehy, and J. H. Schaffner, "Wave-Coupled LiNbO₃ Directional Coupler Modulator at 94 GHz," Paper 4.3, Photonic Systems for Antenna Applications-IV, Monterey CA, 18-21 January 1994.
- 3.4 H.-G. Unger, Planar Optical Waveguides and Fibers, Oxford University Press, 1977, Chapter 2.
- 3.5 S. Thaniyavarn, "Modified 1 x 2 Directional Coupler Waveguide Modulator," Electronics Letters, vol. 22, No. 18, 28 August 1986, p.p. 941-942.

4. EFFECT OF VELOCITY MISMATCH ON MODULATOR LINEARIZATION

4.1 BACKGROUND

Methods of linearizing electro-optic modulators to achieve greater dynamic range have been known for years. A number of these methods were studied at Hughes Research Laboratories under a contract with Rome Laboratories [Ref. 4.1]. Further work led to a summary publication by Bridges and Schaffner, comparing several different linearization schemes for both Mach-Zehnder-based and directional-coupler-based modulators [Ref. 4.2]. However, neither of these publications considered the effects of operating frequency on the degree of linearization, notably the effect of velocity mismatch, wherein the modulating signal propagating along the modulator's electrical transmission line travels slower than the optical signal propagating in the optical waveguide.

The mismatch in velocities limits the operating bandwidth of a simple modulator for a given length, or requires a shorter length (and hence less sensitivity) for a given bandwidth. The question then becomes, does this mismatch affect the distortion products in the various linearization schemes in the same way as they affect the operating bandwidth? If they do, then the dynamic range should be independent of frequency, since both signal and distortion products are reduced by the velocity mismatch. However, if the distortion product reduction sought by the particular linearization scheme is affected differently by the mismatch than the simple reduction in sensitivity, then the dynamic range will not be independent of frequency. Farwell, in his thesis [Ref. 4.3], describes the results of calculations he made taking into account velocity mismatch, and states that the mismatch has a degrading effect with frequency on the dynamic range of his linearized modulator (a directional coupler followed by two dc biased sections).

We wished to extend Farwell's calculations to other linearized modulator schemes, and make the same kind of comparison that Ref. 4.2 did, but now accounting for the effects of velocity mismatch. We accomplished this during the present contract, and published our results in the Journal of Lightwave Technology [Ref. 4.4]. Since that publication, we have continued to look at other effects and other linearization schemes, obtaining results that are not yet published. Accordingly, Ref. 4.4 is included as Appendix A, so that the equations and data therein need not

be repeated in the body of this report. We will give a brief summary of the results contained in that paper below, followed by the results and conclusions of subsequent calculations.

4.2 SUMMARY OF APPENDIX A: J.O.L.T PAPER “BANDWIDTH OF LINEARIZED ELECTRO-OPTIC MODULATORS”

In this section we summarize the general conclusions of Ref. 4.4, included in this report as Appendix A.

The general technique of accounting for velocity mismatch is essentially the same as that used by Farwell [Ref. 4.3]. In a numerical calculation of the type used for analyzing linearization schemes the modulation signal and optical carrier paths through the modulator are now broken into many short segments. The contribution to the modulation by a short segment is calculated using a transfer function that ignores any effects of frequency, including velocity mismatch (as in Ref 4.2, for example). However, in the next segment, the phase of the modulating voltage is shifted by an amount that is appropriate to the mismatch in velocities between the electrical and the optical guides. This phase shift is a function of a universal scaling parameter

$$f \Delta n \Delta z, \quad [4.1]$$

that is, the product of the modulation frequency f , the difference between the effective refractive indices of the modulating and optical waves Δn , and the length of the section Δz .

The net modulation at the end of the modulator of length L is then obtained by adding up the modulation contributions of each of the small segments. The overall result is found to scale simply as

$$f \Delta n L = \text{a constant} \quad [4.2]$$

which is convenient, since the results of any calculation can then be scaled to new values of f or Δn or L . As a test case, the calculations were also run with Δn set equal to 0 for the perfectly velocity matched case to see if the results in Ref. 4.2 (where velocity mismatch was simply ignored) were obtained. The results agreed.

The calculated dynamic ranges vs frequency are given in detail in the curves in Appendix A for several modulators:

- * Dual parallel Mach-Zehnder modulators (DPMZM)
- * Dual series Mach-Zehnder modulators (DSMZM) [denoted "DSMZM" in Ref 4.2 and Appendix A]
- * Simple directional coupler modulator biased at 0.79 V_s (DCM.79) [denoted "SDCM" for "sub-octave DCM" in Ref. 4.2 and Appendix A].
- * Directional coupler modulator with two dc bias sections (DCM2DC) [denoted "DCM2P" in Ref. 4.2 and Appendix A].

The DPMZM, DCM.79, and DCM2DC were analyzed in Ref. 4.2 without velocity mismatch. The DCM2DC is Farwell's modulator and is described in Ref. 4.3, both with and without velocity mismatch. The DSMZM was not treated in Ref. 4.2. It is the modulator developed by Betts at MIT Lincoln Laboratories [Ref. 4.5].

The general conclusions reached in Ref. 4.4 (Appendix A) can be stated simply:

- (1) The effects of velocity mismatch scale as $f \Delta n L$ (as previously stated).
- (2) With the exception of the DPMZ and a special case of the SDSMZ, all linearized modulators exhibit degradation in dynamic range with frequency. If the dynamic range is optimized at zero frequency, then there is a roll-off in the dynamic range at higher frequencies. By readjusting the exact bias conditions, the original dynamic range can be recovered at some finite frequency, but then there will be a roll-off on both sides, giving a band-pass characteristic to the dynamic range.
- (3) Some modulators exhibit much faster roll-off in dynamic range than others. For example, the DCM2DC rolls off about two orders of magnitude faster with frequency than the DCM.79.
- (4) Some modulators that can be adjusted for zero second harmonic when perfectly velocity matched (e.g., the simple DCM.43) exhibit significant second harmonic at higher frequencies,

and can have a dynamic range limited by the second harmonic rather than intermodulation distortion above that frequency. Other modulators (e.g., the DPMZM) retain their second harmonic-free behavior at all frequencies.

(5) The DPMZM exhibits a dynamic range independent of velocity mismatch. Even though the sensitivity of the modulation falls off as a sinc function of the velocity mismatch, both signal modulation and intermodulation distortion are affected equally. And the phase shifts through the modulator for the two intermodulation components generated in the two MZM's that must cancel are guaranteed to be equal by symmetry. Thus, the cancellation process is independent of frequency. This may be contrasted with the general case of the SDSMZ, in which the two intermodulation components generated in the two MZM's traverse paths with different phase shifts, so that the cancellation (and hence the dynamic range) becomes a function of frequency. The exception occurs when the spacing between the two MZM's is such that the optical transit time between the two MZM's takes an integral number of cycles at the modulation frequency [see (6) below].

(6) The roll-off in dynamic range may be countered in some modulators by segmenting the drive electrodes, and re-phasing the drive at the beginning of each electrode. (This is "velocity matching on the average", used by the antenna array modulators of Sections 3 and 5; it restores the linearization as well as the signal sensitivity). In some modulators (e.g., the DCM.43 and DCM.79) the roll-off is monotonically improved with segmentation. In others (e.g., the SDSMZM) the roll-off is monotonically improved with 3,5,7... segments, but is eliminated completely with even numbers of segments 2,4,6,... (This is the "special case" of the SDSMZM that has dynamic range independent of frequency, stated in item (5) above.)

The difference between MZM-based linearized modulators and DCM-based linearized modulators cited in (1) and (2) above may seem a bit puzzling at first, since the simple MZM and the simple DCM perform so much alike otherwise. In fact, even the effects of velocity mismatch on the signal are much the same for the simple MZM and simple DCM, with the simple DCM actually exhibiting a slightly larger signal operating bandwidth limited by velocity mismatch. Why the notable difference in the two kinds of linearized modulators?

A simple argument goes as follows: All linearization schemes are essentially “cancellation” schemes wherein different ratios of signal-to-intermodulation distortion are generated in different pathways, then scaled and combined so that the intermodulation distortion products from the two pathways cancel exactly, while the signals cancel only partially. This is most easily seen in the DPMZ scheme, where the two pathways and the two modulation processes are clearly identical, and the scaling is accomplished by the optical split and the modulation voltage split between the two modulators. And in the DPMZ scheme, one can also see that the effect of velocity mismatch degradation should be the same on both signal and intermodulation, so that no additional effects should be felt on the linearization. While the signal (and intermodulation) roll off with frequency, the dynamic range is independent of frequency.

The argument is a little less clear-cut for the DSMZM, since there is a single pathway through both modulators. The two modulators may be identical but the pathways are not. Both the r-f voltage and the optical signal from the first modulator travel an additional (velocity mismatched) path through the second modulator. However, if the modulating voltage supplied to the second modulator is re-phase to account for this differential delay from the additional path, then the two paths are essentially identical, and no additional degradation occurs in the linearization. Of course, both signal and intermodulation still roll off due to the velocity mismatch within each MZM, but the dynamic range is independent of frequency.

For the directional-coupler-based modulators treated in Ref. 4.4, the situation is less clear. The creation and combination of signals and intermodulation products in modulators containing more than one directional coupler are not localized. The creation and combination process is a function of transit time through the device and turns out to be typically much more sensitive to timing errors than simple signal degradation. However, a directional coupler modulator not treated in Ref. 4.4, the Y-Fed DCM proves an interesting exception to the general conclusions reached in Ref. 4.4. Because of its inherent symmetry, its linearization proves to be robust with frequency. Although the signal and intermodulation generation and combination processes are still distributed, the symmetry of the device makes them identical. We have analyzed this modulator in detail, and the results are presented in Section 4.5.

4.3 SUMMARY OF APPENDIX B: U.S. PATENT 5,886,807, TRAVELING-WAVE REFLECTIVE ELECTRO-OPTIC MODULATOR

In the course of studying the dual series MZM, we found that the linearization (dynamic range) could be made independent of frequency if the rf drive to the second MZM could be rephased to match the extra delay in the optical path between the first and second modulator. In principle, this could be accomplished in two ways: (1) Speed up the electrical wave reaching the entrance of the second modulator so that it is in the proper phase to match the modulation already on the optical signal from the first modulator, or (2) Slow down the optical wave between the two modulators to accomplish the same thing. Most schemes of true velocity matching focus on solution (1). Solution (2) is usually discarded, since it would seem quite difficult to add extra transit time to the optical path; it is typical waveguides used in lithium niobate modulators are only weakly guiding, and would have extremely large radiation losses for any bends that could be put on a chip. One might imagine a folded path using several turning mirrors, but such a solution also seems impractical to implement.

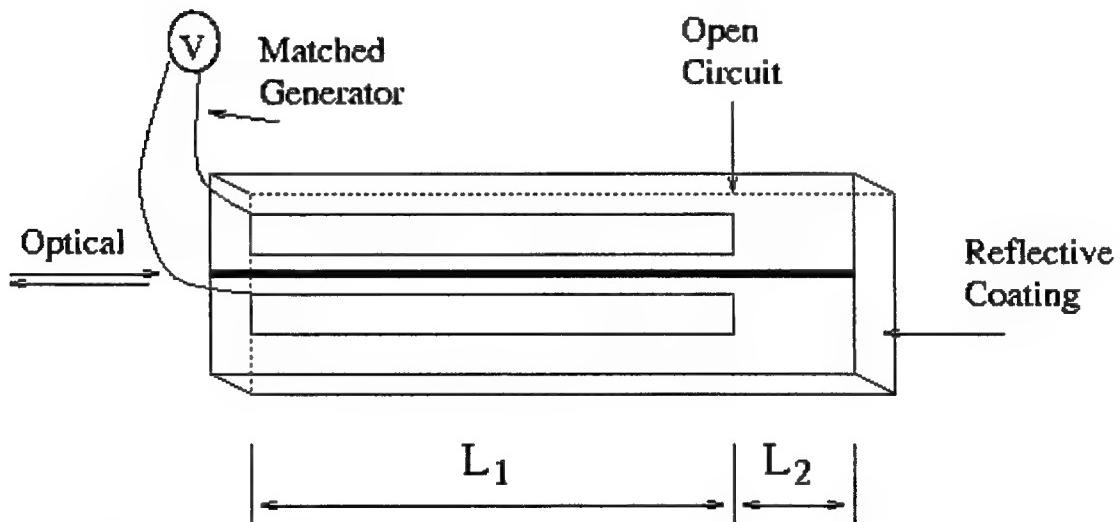


Figure 4.1 A reflective traveling wave phase modulator.

However using a single reflection, namely, a retro-reflection, can accomplish everything we need, and that is the crux of the invention in Ref. 4.6. Consider the reflective phase modulator

shown in Fig. 4.1 (The phase modulator is used as a simple example, but the invention works just as well with a MZM). The optical signal enters at the left and traverses the electrode region L_1 , where it is modulated. The light then traverses the additional distance L_2 and is reflected by a mirror deposited on the end of the modulator. (An external mirror or pigtailed fiber extension could also be used, but this would produce additional losses at the various additional interfaces). The r-f signal traveling down the distance L_1 is reflected at the end by the open-circuit nature of the transmission line. This reflection is of the same magnitude and phase as the incident wave. Thus the optical wave and the r-f wave are again traveling in the same direction back to the input of the modulator, so that a second pass of phase modulation occurs. However the light has traversed a distance $2L_2$ compared to the r-f. If we make the distance L_2 that given in Eq. 4.3,

$$L_2 = 0.5[(n_m/n_o) - 1] L_1 \quad [4.3]$$

then the optical wave and r-f wave will be in phase during the reflected path, and the dynamic range will be independent of frequency. For LiNbO₃, $n_m = 4.0$ and $n_o = 2.2$. Thus $L_2=0.41 \times L_1$. The modulator must extend 41% past the end of the r-f electrode. This extra real estate can be usefully employed. Most modulators have at least one bias electrode. This may be the r-f electrode itself, or an additional electrode with only d-c applied. Such an additional electrode can be placed between the end of the r-f electrode and the reflective end of the modulator.

This modulator is now equivalent to a dual-section phase modulator with length $2L_1$ with the electrical and optical wave re-phased at the second segment. To do this in a dual section phase modulator, the r-f power would have to be split into two parts, one driving each section. However, since the reflective modulator of Fig. 4.1 does not have to split the r-f power, but uses the same power twice (first as a forward-travelling wave, then as a backward traveling wave), the electrical voltage wave is 1.44 larger than it would be in the modulator with two physical segments and thus has twice the gain (+3 dB) as the analogous two section phase modulator. It also uses only half the crystal real estate. Of course, it does have the disadvantage that an optical circulator must be used to separate the input from the output of the modulator.

This same optical-delay-plus-reflection scheme can be used with other common modulator configurations. The Mach Zehnder modulator has similar attributes to the phase modulator of

Figure 4.1. Thus the dual series MZM can be implemented as a single MZM with reflection. And for this configuration of the SDSMZM, the dynamic range is independent of frequency.

The directional coupler based modulators can also be used with this kind of re-phasing scheme. They also have an added feature: the designer can choose the electrical length to be either 0.5π or π . If a DCM is used with a mirror covering both ports at the right hand edge, then the reflected output can be taken either at the original left hand input port (π length) or at the other left hand input port (0.5π length). Depending on the system application, one or the other may be desirable. In a remote application, it may be desirable to have only one fiber and not to have the laser near the modulator. In this case the designer would choose the former and use an optical circulator to separate the outgoing and incoming signals on the fiber. In other applications the designer could choose the 0.5π with separate input and output ports, which would not require an optical isolator.

In the case of DCM-based linearized modulators, the dynamic range is not independent of frequency, as we've seen in Section 4.2 and Appendix A. However, the roll-off may be decreased by rephasing, and the reflection scheme should yield the improvement indicated by the two-section rephasing curves given in Appendix A.

4.4 ODD-LENGTH DIRECTIONAL COUPLER MODULATORS

The frequency-dependent gain of the directional coupler modulator is a mathematically rich function, unlike the simple gain function of a Mach-Zehnder. Unfortunately, it is also not solved in analytic terms. Using the numerical calculation method described in Ref. 4.4 and Appendix A, we have discovered interesting properties of the gain of a standard directional coupler modulator. Some of these interesting features are presented below. There are likely even more interesting modes of operation and variations on this theme that we have not had time to explore fully.

The gain function of a Mach-Zehnder modulator can be expressed functionally as

$$G_{MZM} = A(V_B, P) \times [\sin(2\pi f \Delta n L) / (2\pi f \Delta n L)]^2 \quad [4.4]$$

Note that the gain can be broken into two factors, the first of which contains the modulator physical parameters (denoted "P", which include the length L, the electro-optic coefficient of LiNbO₃, etc.) and the bias voltage, but not the frequency, and a second [sinc]² term that contains the frequency, but not the bias voltage. Thus the frequency behavior is quite simple and explicit.

In the case of the directional coupler modulator, this separation of functional dependence of the frequency from the rest of the modulator parameters is not possible. Thus the bias voltage and all the other physical parameters affect the frequency dependence. Since no reasonable analytic expression exists from which the general behavior with frequency can be seen, we undertook to explore this functional dependence numerically with the program we had developed.

Normally, a simple DCM is fabricated with an electrical length of 0.5π , which produces a complete cross-over in the optical signal from the input optical waveguide to the other optical waveguide. Then the modulator is biased to minimize second harmonic generation (i.e., zero second derivative) which is 0.43 of the value required to transfer the optical signal back to the original input optical waveguide. This length and bias point produce equal modulated outputs of half the light in each of the output waveguides. This choice of bias is also close to the value that produces the maximum link gain, i.e., maximum first derivative. (In the MZM, the bias of 0.5 V_π simultaneously zeros the second derivative and maximizes the first derivative, but in the DCM 0.43 V_S does not quite maximize the first derivative, but it is quite close.)

We wished to explore the gain variation with frequency of DCM's of different electrical lengths, not just 0.5π . But since the frequency variation also depends on bias voltage, we had to choose a bias condition that would make sense for all electrical lengths. Unfortunately, the condition of "zero second derivative" or zero second harmonic generation doesn't make sense. So instead, we chose the bias voltage that maximizes the gain at zero frequency. Figure 4.2 compares the link gain variation with frequency for DCM's of varying electrical length, but all odd half multiples

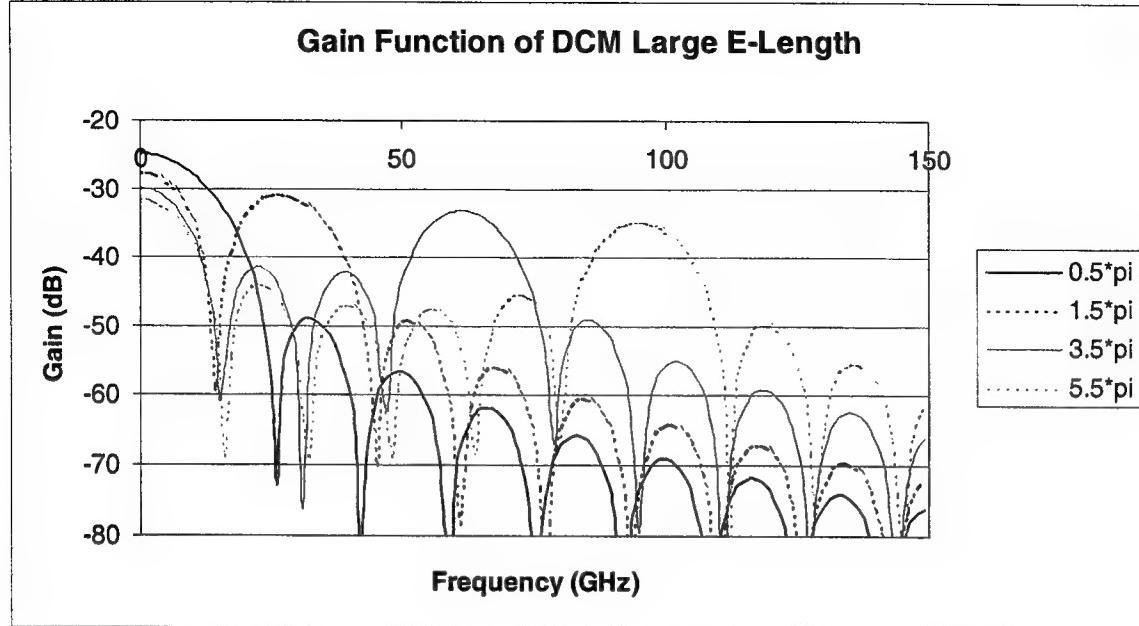


Figure 4.2. Gain of a link with a simple DCM using the canonical parameters of Refs. 4.2, 4.4 and Appendix A.

of π . (The other link parameters are chosen as the “canonical” parameters used in Refs. 4.2, 4.3 and Appendix A, so that comparisons with those works can be made). This family of curves may seem a little confusing at first, but the essential features are easily extracted. The electrical length 0.5π produces the highest link gain at zero frequency (-24.8 dB), which then falls off to a null at about 26 GHz. It then recovers to a value of -54 dB at 37 GHz, and drops to another null, and so on. An electrical length of 1.5π produces a slightly lower overall gain at zero frequency, -27.5 dB, and drops sharply to a null at a lower frequency of 16.7 GHz (actually, the same null that a MZM of the same parameters would have!). The gain then recovers greatly to peak at about -31 dB at about 25 GHz before dropping to the next null at about 49 GHz, following which it also varies aperiodically with frequency, but never comes close to the behavior at zero frequency or 25 GHz. A modulator of length 3.5π (we have omitted 2.5π and 4.5π curves to simplify the plot) has a zero frequency gain of -30 dB, then drops drastically to a null at 16.7 GHz (the same frequency as the 1.5π case). From this null, it recovers only a little, rising to a peak of -42 dB at about 20 GHz, followed by another null at 33 GHz, and another peak of -43 dB at about 40 GHz, then falling to another null at 49 GHz. But then it rises to a maximum of -33 dB around 65 GHz! Then it falls off to a null at 80 GHz, and decays in a further set of much

smaller peaks and nulls. This curious behavior is repeated for longer modulators. For a 5.5π long modulator, zero frequency gain is about -32 dB, followed by three peaks and 4 nulls, before a peak of about -34 dB is reached at about 90 GHz. Note that this peak is only about 2 dB worse than the zero-frequency gain! And this for a fully velocity mismatched ($\Delta n = 1.8$) modulator. The message here would seem to be that if only a band-pass link is required, a 90 GHz modulator could be made with perhaps 15 to 20 GHz bandwidth and only 7 to 10 dB loss from a perfectly velocity matched 0.5π length DCM link!

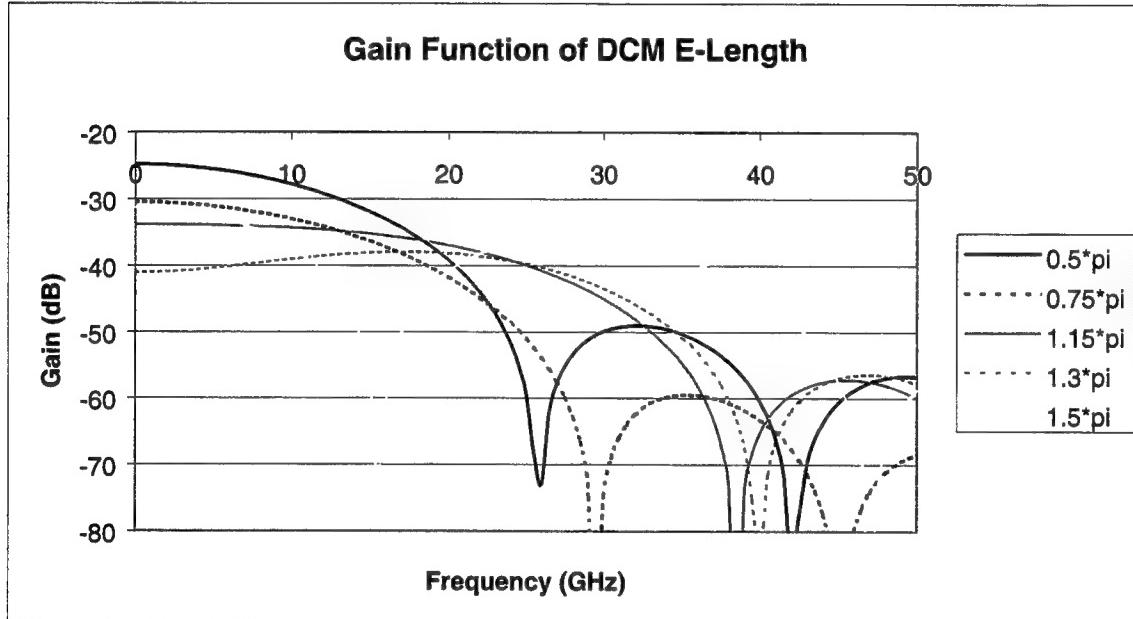


Figure 4.3 Gain for a link using a simple DCM with the parameters given in Refs. 4.2, 4.4, and Appendix A

We also explored the range of DCM electrical lengths other than odd half multiples. Figure 4.3 shows the link gain for some lengths less than 1.5π .

Again, the bias voltages are chosen to maximize the zero-frequency gain (which is close to the second harmonic null value). The curve for 0.5π again shows a maximum value of -24.8 , falling to a null at 26 GHz, and recovering to a peak of about -49 dB at about 32 GHz, as in Fig. 4.2. The 1.5π curve is also repeated, showing a null at 16.7 GHz and a peak of -36 dB at 28 GHz.

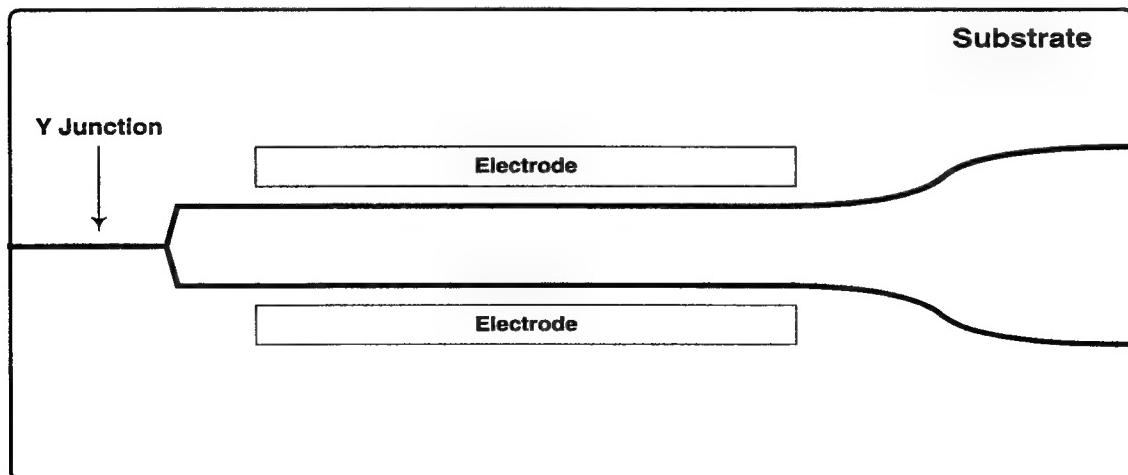
But the other lengths exhibit interesting behavior, and a few examples are shown. By changing the length, the location of the first null can be greatly moved out in frequency, albeit at a sacrifice in zero-frequency gain. A length of 1.15π moves the first null to 38.5 GHz, and produces a response that has a -3dB bandwidth of 20 GHz, twice as much as the -3dB bandwidth of the 0.5π length modulator. However, this is obtained at a sacrifice of 9dB in zero-frequency gain, and only a 2 dB gain improvement over the 0.5π -length modulator at 20 GHz. A choice of 1.3π for the length moves the first null out to 40 GHz, and flattens the low-pass bandwidth even further, with a “ripple” response. The -3 dB bandwidth of this link is about 27 GHz.

There are many more calculations that could be made. For example, it would be interesting to repeat these calculations for linearized DCM modulators, such as the DCM.79 or the DCM2DC. Unfortunately, there was no time remaining in the present contract to do so. There may be some surprises still lurking in the DCM transfer function.

4.5 THE Y-FED DIRECTIONAL COUPLER MODULATOR

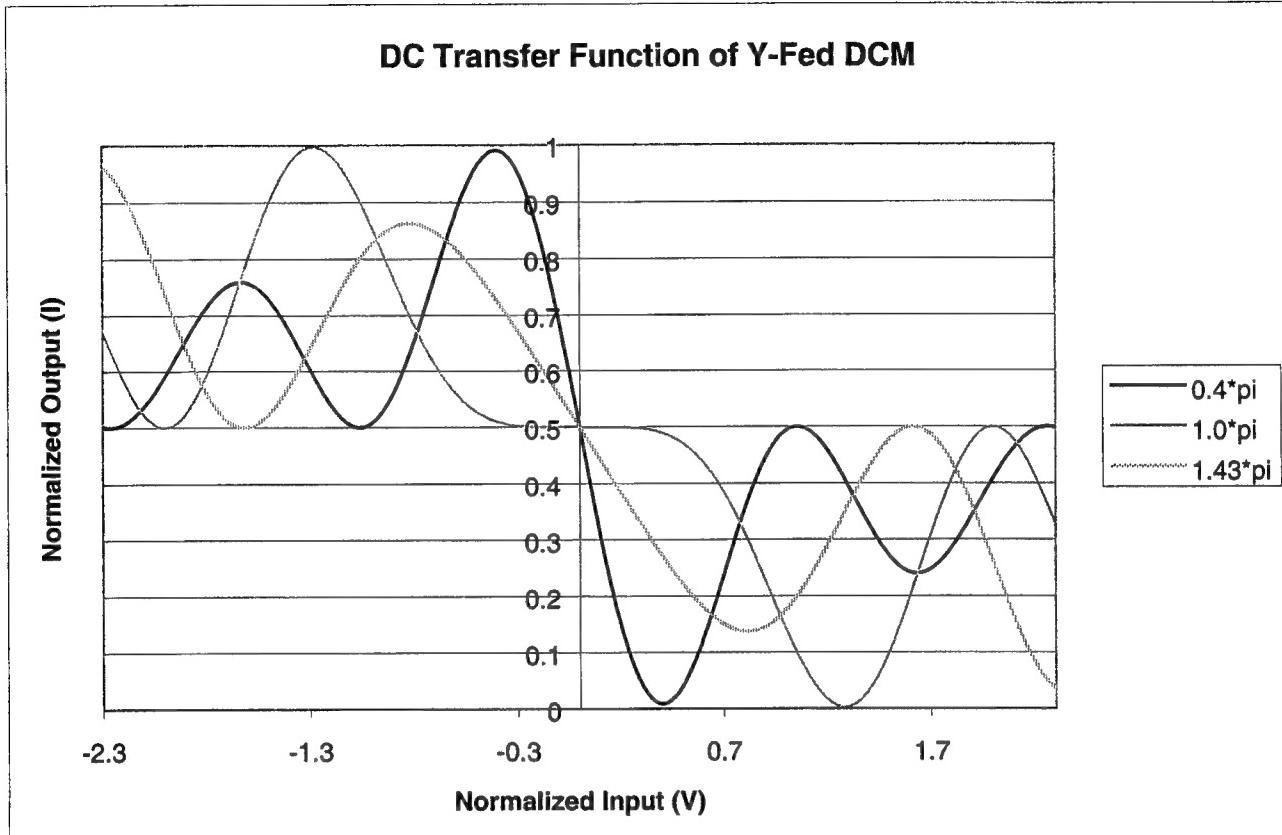
Near the end of this contract, we looked at a still another DCM based modulator, the so-called Y-fed directional coupler modulator (YFDCM). This modulator, originally reported several years ago, by Thaniyavarn [Ref 4.7], was used more recently in a linearization scheme by Tavlykaev and Ramaswamy [Ref. 4.8]. We wished to see how robust this scheme was with velocity mismatch. (Ref 4.8 does not include frequency effects). In doing so, we found some very

Figure 4.4 Schematic representation of a YFDCM.



interesting behavior of this modulator, in addition to understanding its behavior with velocity mismatch.

Figure 4.5 Transfer function of a YFDCM



The Y-fed directional coupler is simply an ordinary DCM in which the optical wave is fed into both arms of the input, in equal amounts. A Y-branch junction is typically used for this (hence the name), but any other means of feeding both arms equally will give the same result. Such a modulator is shown schematically in Fig 4.4. The important effect that this method of excitation has on the device transfer function is to make that function anti-symmetrical about the point (0, 0.5), as shown in Fig. 4. 5. This means that with zero bias applied to the electrodes, there are no even-order harmonics, due to the excitation symmetry alone. Recall that the ordinary simple DCM requires a bias of $0.43 V_s$ to make the second harmonic zero, but all other even-order harmonics are non-zero at this value of bias. And no bias stabilization circuitry is required for the YFDCM since the proper bias is exactly zero. Thus the YFDCM looks better at the outset than a simple DCM.

In Ref. 4.8, Tavlykaev and Ramaswamy show that the third derivative may also be made zero at zero bias by choosing a specific electrical length, 1.43π , for the modulated directional coupler section. This choice of length reduces the intermodulation distortion and produces an increase in dynamic range. The improvement is about the same as any other linearization scheme.

Figure 4.6 Signal and third-order intermodulation products for a YFDCM with parameters the same as modulators studied in Ref. 4.2. Velocity mismatching has been ignored. The electrical length is 1.43π .

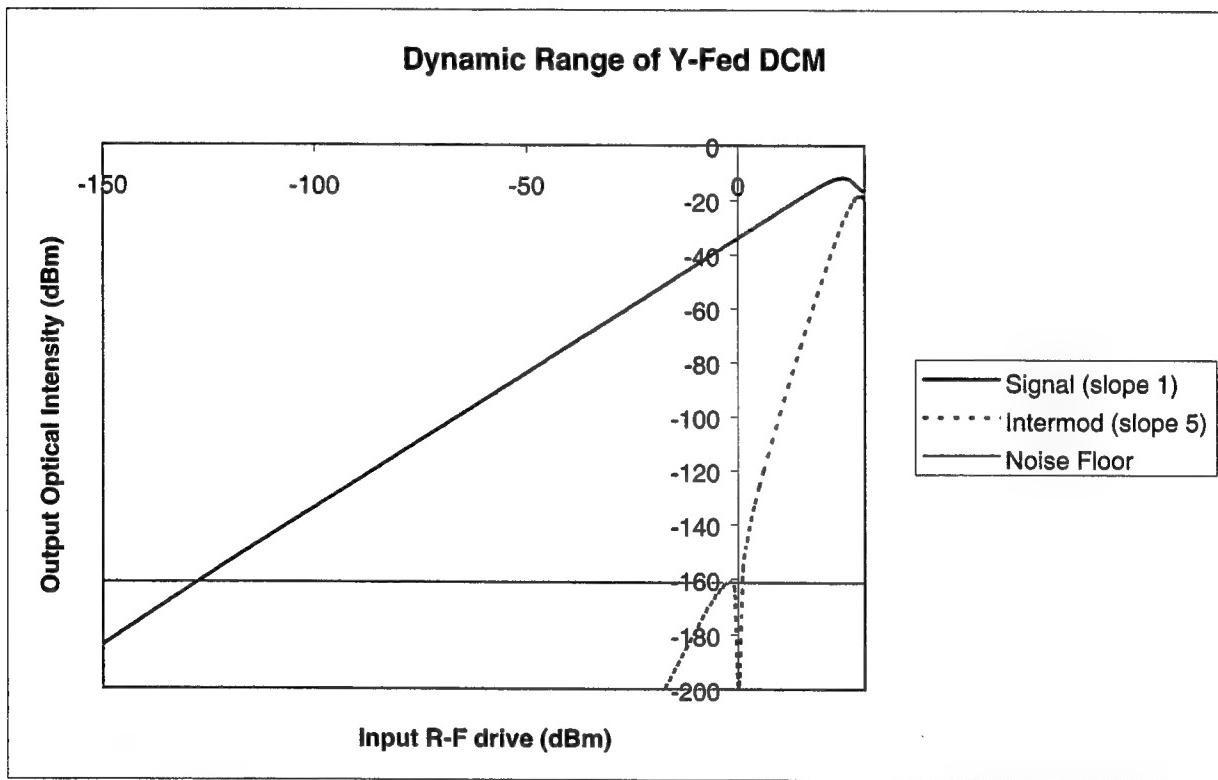


Figure 4.6 shows the signal and distortion for a Y-fed DCM with the canonical parameters of Ref 4.2, neglecting the effects of velocity mismatch. The dynamic range is about 125 dB, as seen from the figure. Our calculations agree with those in Ref 4.8

We also looked at the sensitivity of the linearized YFDCM to its physical parameters. Since the bias is zero, there is no sensitivity to bias voltage. However, the modulator is quite sensitive to the choice of electrical length. Figure 4.7 shows this sensitivity for a 1 Hz bandwidth. This

sensitivity may be more of a problem than the sensitivity of other linearized modulators to bias voltage error or power-splitting error, since this is a “fabricational” sensitivity. There is no

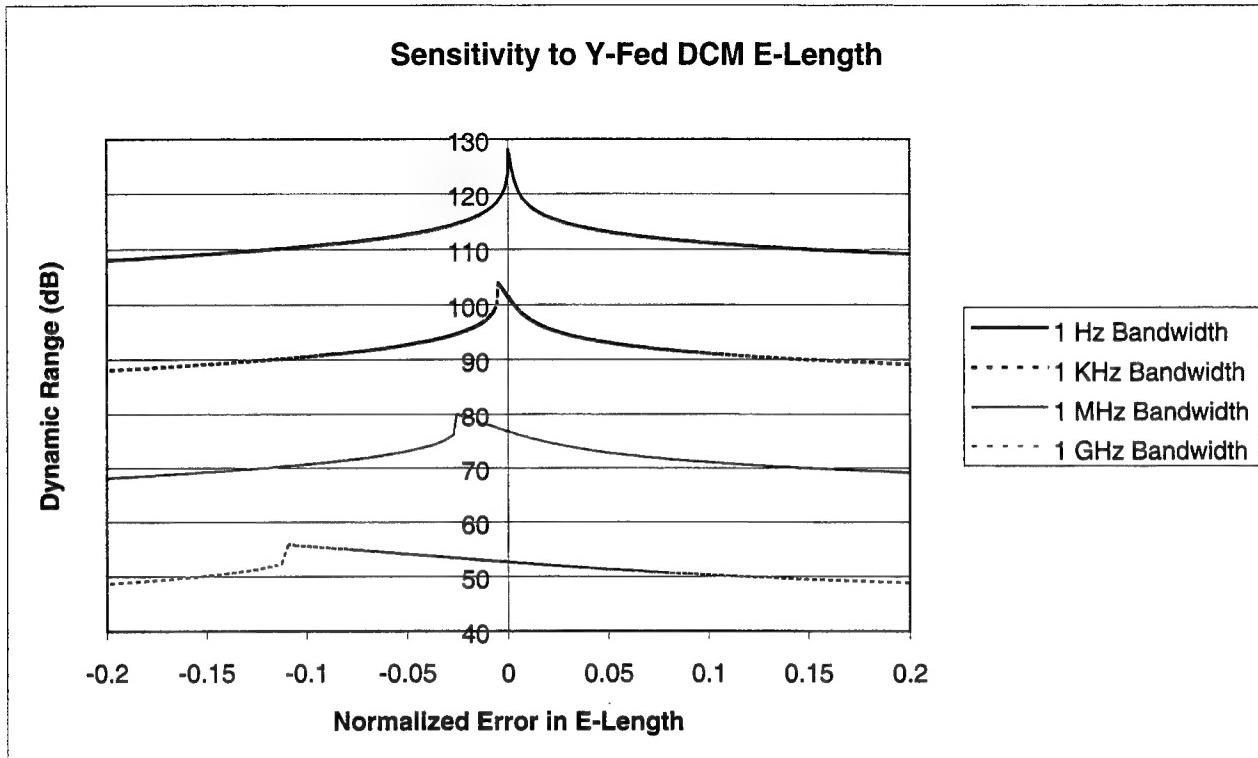


Figure 4.7 Sensitivity of dynamic range to modulator electrical length for a linearized YFDCM using the parameters of Ref. 4.2. Velocity mismatch is neglected.

“knob-on-the-power-supply” to turn to adjust it. The DCM active length must be fabricated to its exact length to this accuracy. Clearly, this cannot be done with a mask alone. We propose that the modulating electrodes be made slightly longer, then etched away or laser-ablated away while monitoring the actual intermodulation amplitude. However, once this trimming is done accurately, there should be no reason for the performance to drift around, except for the processes described in Chapter 6 of this report.

We also looked at the YFDCM as a function of frequency, assuming the maximum velocity mismatch, for a number of different active lengths. This was actually a simpler task than for the ordinary DCM, since bias voltage is no longer a variable; it is always zero for the YFDCM.

Figure 4.8 shows the results for electrical lengths of 0.5π , 1.5π , 3.5π , and 5.5π . The standard value is 0.5π for a simple DCM. The curve for the electrical length for which the modulator is linearized, 1.43π , is nearly overlapping with 1.5π , so it need not be drawn. Note that the gain of the 0.5π length YFDCM drops off with a null at 26 GHz, then recovers to a value 23 dB below its zero frequency value, then another null, etc., very similar to standard DCM. For an electrical length of 1.5π , the zero frequency gain drops to -34 dB, then nulls at 8.5 GHz, but then recovers to a value of -26 dB, only 1.6 dB worse than the zero frequency gain for the 0.5π modulator! Then it nulls again at 42 GHz, and so on. Longer electrical lengths show a band-pass response pushed out to still higher

Again, referring to Fig. 4.8, note that the curve for a length of 1.5π (very close to the curve for 1.43π , which was not plotted), is also a band pass modulator. The response drops off rapidly with frequency from zero frequency, then recovers to peak around 22 GHz. To be able to use the linearized YFDCM of Ref. 4.8 as a low-pass modulator, Fig. 4.8 shows that very careful velocity matching must be used, to extend the bandwidth to more than the 4 GHz that you would obtain with a velocity mismatch $f \Delta n L = 1.8$.

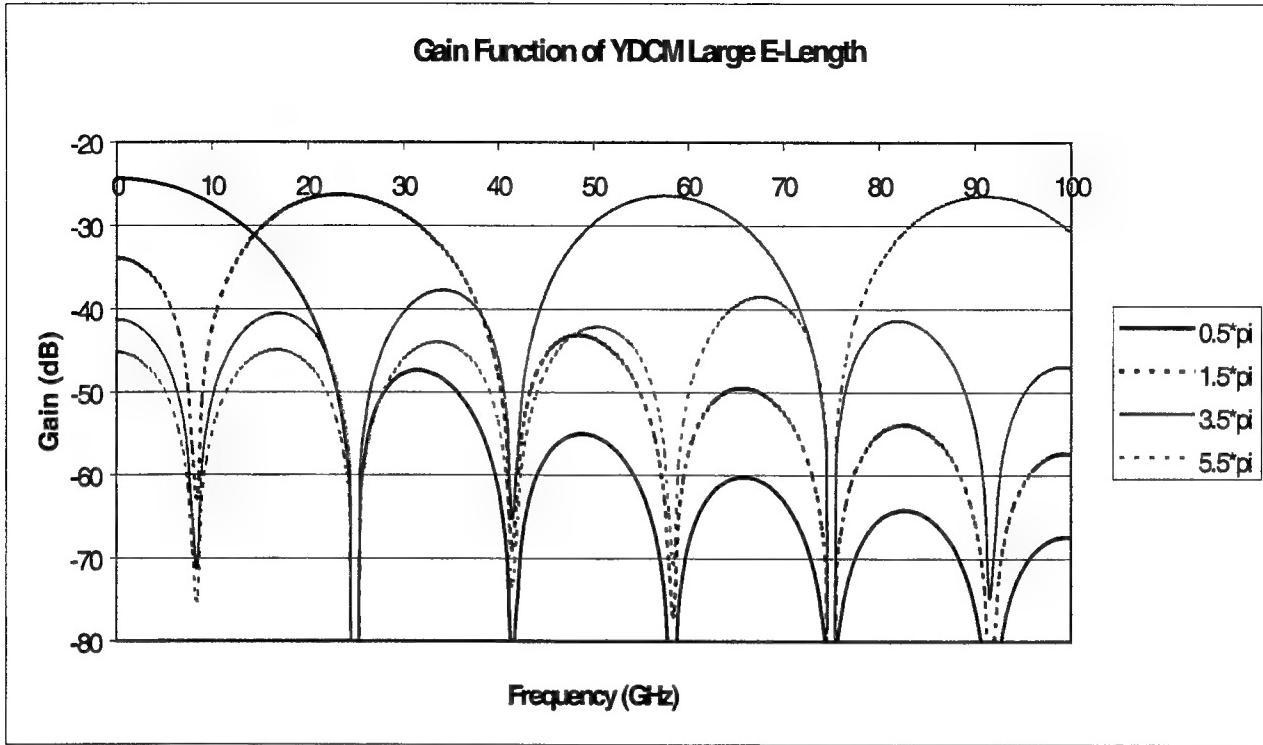


Figure 4.8: Gain of Y-Fed DCM using the parameters of Ref. 4.2

We have derived a simple first order approximation of the gain function for YFDCM's [Ref. 4.9]:

$$G_{YFDCM} \approx A(V_B, P) [\text{sinc}(\gamma - \theta)]^2 \quad [4.5]$$

where γ is given by

$$\gamma = 2\pi f \Delta n L \quad [4.6]$$

and θ is the electrical length of the directional coupler, in radians.

Because of the physical symmetry of the YFDCM, the frequency dependence of the gain can be extracted in the sinc-function form shown in the equation, similar to Eq. 4.4. This expression is most accurate around the center frequency of the passband; it gives the location and width of the passband accurately, and almost gives the precise shape. It is sufficiently accurate that the differences from the full numerical method would not likely be noticeable in an experimental measurement. It is inaccurate in the sidelobes, however, especially many sidelobes away from

the passband. We have derived another approximate expression [Ref. 4.9] that contains a multiplicative correction for the variation with frequency in the sidelobes,

$$G_{YFDCM} \approx A(V_B, P) \exp[-2(\gamma - \theta)/\theta] [\text{sinc}(\gamma - \theta)]^2 \quad [4.7]$$

Figure 4.9 shows a Y-Fed directional coupler with electrical lengths between 0.5π and 1.5π also at zero bias. Note that the peak gain is actually achieved for $\theta=0.4\pi$, not the standard length of $\theta=0.5\pi$. Additionally at this length, as well as some others, the gain does not null completely,

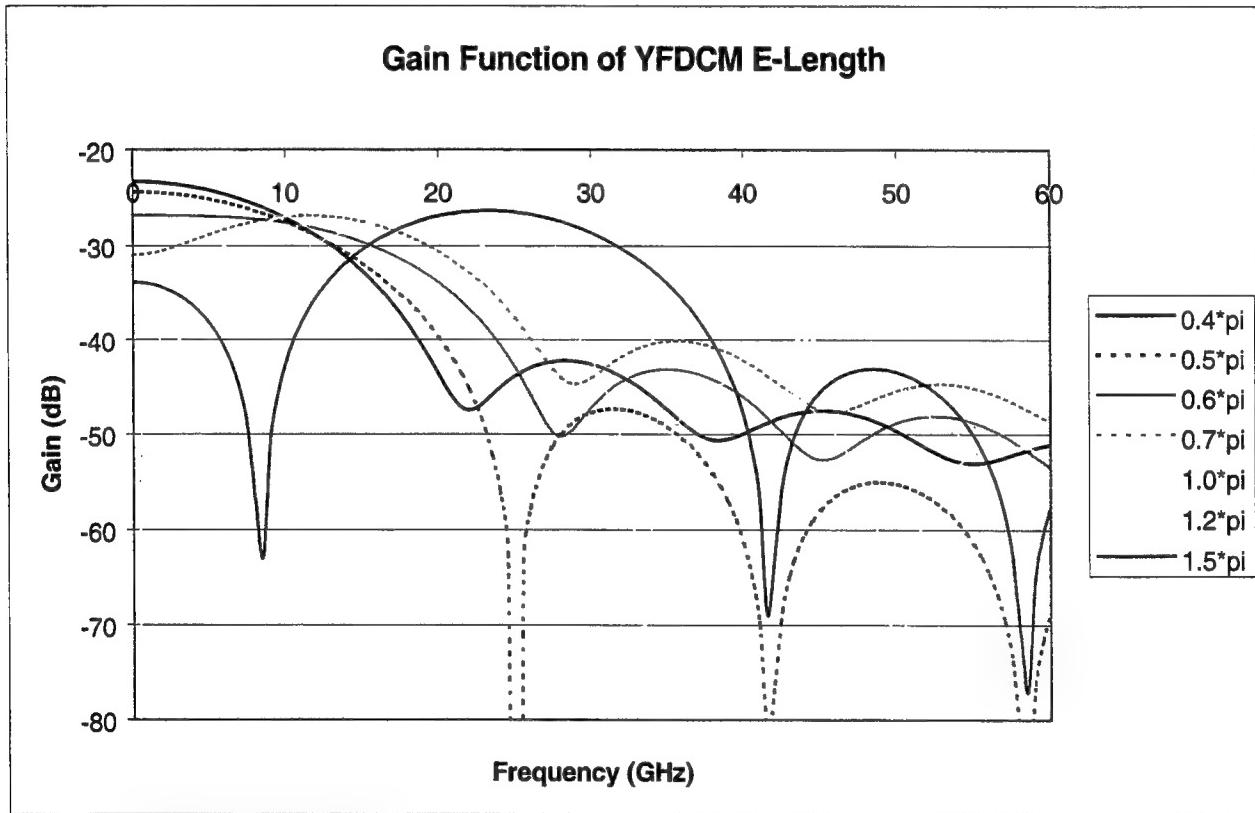


Figure 4.9 Gain of Y-Fed DCM below $\theta = 1.5 \pi$

although it does fall by at least 25 dB. There are two optical paths in a directional coupler, and if their null positions do not exactly coincide there will be residual modulation. Also note the position of the nulls. A standard Mach-Zehnder with equivalent parameters nulls at 16.7 GHz. The standard Y-Fed DCM nulls at 25 GHz, an improvement of almost 40% over the Mach-Zehnder at the 3-dB point. This is a “one time” effect; the nulls at higher frequencies than the

first are separated by about 16 GHz, resulting from the underlying γ of the modulator. The 1.5 curve exhibits this passband property, as do all higher values of θ .

The electrical length may be chosen to make a flat modulator response over a wide band even in the presence of loss. The 0.6π curve is flat, losing only 1-dB out to 10 GHz, unlike the response for an electrical length of 0.5π or the response for a Mach-Zehnder. An electrical length of 0.6π - 1.0π demonstrates a variable amount of peaking. The zero-frequency value may be 3-10 dB below the optimum, but some of this is regained, so that at around 10-15 GHz the response is only a couple of dB below the zero-frequency value. Given this novel property and a known value for frequency-dependent conductor losses, the electrical length may be chosen to compensate for the losses providing a flat response over a wide band. Equalizing the response of a traditional link can require complex and expensive hardware. There is a genuine opportunity with this modulator to make inexpensively and compactly a broadband link with a flat frequency response.

It is important to note that Figs. 4.8 and 4.9 show the frequency response of the modulator gain, not the dynamic range. This research program ended before we were able to make the calculations for the frequency dependence of the dynamic range for the YFDCM. We anticipate that the bandwidths for the dynamic range will be significantly smaller than those for the gain alone. This will be true for the linearized modulator scheme of Ref. 4.8 as well. Thus it will be important to make these calculations in the future. The results will be included in Ref. 4.9 and future publications.

4.6 REFERENCES

- 4.1 J. Schaffner, W. Bridges, C. Gaeta, R. Hayes, G. Tangonan, R. Joyce, and J. Lewis, "High Fidelity Microwave Remoting," Hughes Research Laboratories Final Report on Contract F30602-91-C-0104, with USAF Rome Laboratories, November 1992.
- 4.2 William B. Bridges and James H. Schaffner, "Distortion in Electro-Optic Modulators," IEEE Trans. Microwave Theory and Techniques, vol. 43, p.p. 2184-2197, September 1995.

4.3 M. L. Farwell and W. S. C. Chang , "Simulating the response of coupled channel and interferometric modulator designs," J. Lightwave Technology, vol. 13, p.p.2059 - 2068, October 1995.

4.4 Uri V. Cummings and William B. Bridges, "Bandwidth of Linearized Electro-Optic Modulators," J. Lightwave Tech., vol. 16, p.p. 1482-1490, August 1998.

4.5 G. E. Betts and F. J. O'Donnell, "Microwave analog optical links using sub-octave linearized modulators," IEEE J. Photonic Technology Letters, vol. 8, p.p. 1273-1275, September 1996.

4.6. Uri V. Cummings, "Traveling-Wave Reflective Electro-Optic Modulator," U. S. Patent 5,886,807, March 23, 1999.

4.7 S. Thaniyavarn, "Modified 1 x 2 Directional Coupler Waveguide Modulator," Electronics Letters, Vol. 22, No. 18, 28 August 1986, p.p. 941 – 942.

4.8 R.F. Tavlykaev and R.V. Ramaswamy, "Highly Linear Y-Fed Directional Coupler Modulator with Low Intermodulation Distortion,", J. Lightwave Technology, Vol 17 No. 2, pages 282-291, February 1999.

4.9 Topics in Electro-optic Modulation, Uri V. Cummings, Ph. D. Thesis, California Institute of Technology, anticipated June 2001.

5. The Slot-Vee Antenna-Coupled Modulator

5.1 Background

The slot-vee modulator was conceived by Finbar T. Sheehy toward the end of his graduate studies at Caltech, and is described in his thesis [Ref. 5.1] and in the final report on our previous Rome Labs contract [Ref. 5.2]. The idea of using slot-vee antennas grew out of Sheehy's earlier studies of dipoles and vee antennas on the surface of the lithium niobate modulator substrate. Such "normal" antennas are strongly directional into the crystal substrate at an angle that depends primarily on the mm-wave refractive index and only secondarily on the geometry of the antenna itself. Sheehy was seeking an antenna that would have a principal lobe directed parallel to the surface of the crystal. In this way, the crystal itself could act as a dielectric waveguide to direct the wave to the antenna array used for coupling to the modulating transmission system. He realized that such a directive antenna could be possible only if there were a conducting plane separating the air half-space from the lithium niobate half-space. This would mean that the antenna elements would then become "slots" in the conducting plane. From his work on vee antennas, he knew that a slot-vee antenna should be capable of a broader band response than simple dipoles, or slot dipoles. It turned out that Alina Mousessian, then a graduate student in Prof. David B. Rutledge's group, was making measurements on just such a slot-vee antenna in the microwave range [Ref. 5.3], so Sheehy had the encouragement of knowing that such antennas actually worked.

The vee and slot-vee antennas are compared in Fig. 5.1. The two antennas are shown as planar electrodes (black) at the interface between two dielectric regions, one with higher dielectric constant (shaded) and one with lower dielectric constant (white). The vee antenna on the left has two wedge-shaped dipole elements fanning out from the feed point. The angle of the fan is chosen to maximize radiation in the direction bisecting the fan angle. The elements are wedge-shaped to broaden the frequency response of the antenna. Because the antenna resides at the dielectric interface, there is a preference for the radiation to go predominantly into the higher dielectric constant region. Even a small ratio between the two dielectric constants

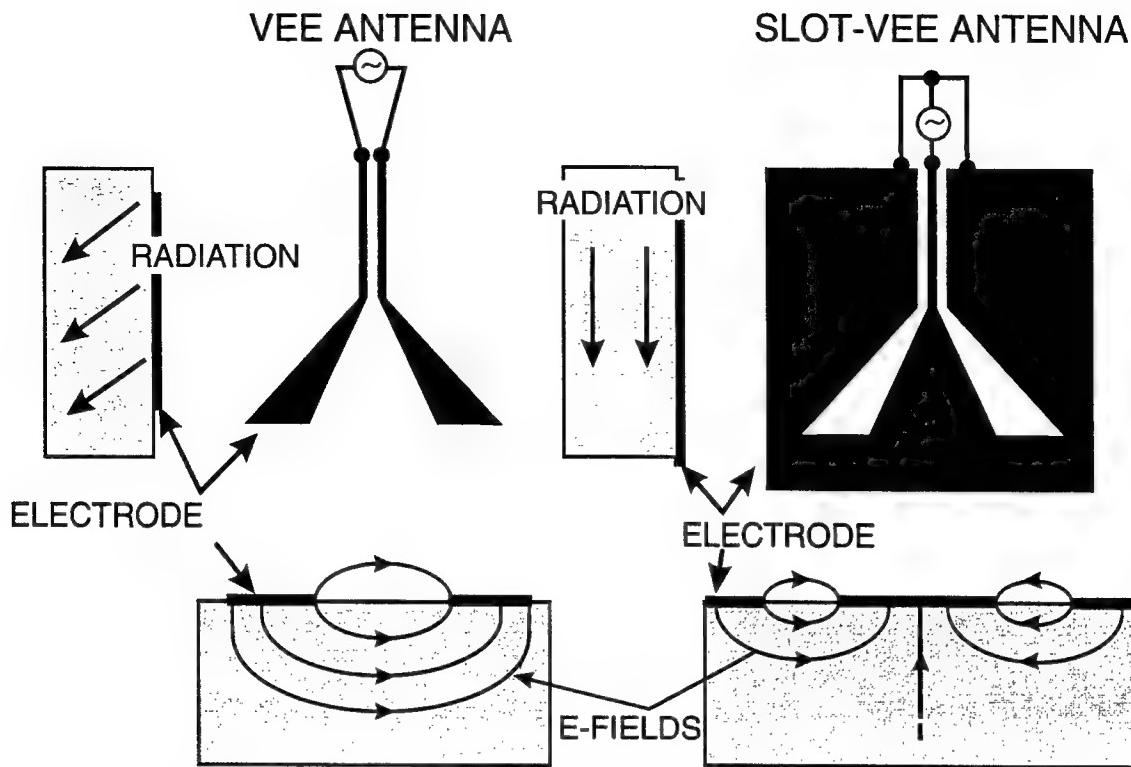


Figure 5.1 Comparison of vee and slot-vee antennas on a dielectric substrate.

(say 4) will put most of the power into the substrate. For lithium niobate and air the ratio is 28 or 41, depending on the polarization of the radiation with respect to the crystal axes. The electric field lines go from one half of the dipole to the other, so that the radiation is polarized normal to a vertical plane bisecting the dipole elements and perpendicular to the dielectric interface. Because of this symmetry, the vertical components of the electric field cancel in the forward direction. The direction of the radiation maximum is also in this plane, downward at an angle of about 24 degrees to the interface normal. Note that the feed line coupled to the antenna feed point is a two-wire (actually, two-strip) line on the dielectric-air interface.

The slot-vee antenna on the right is the complement of the vee. Open space is replaced by conductor and conductor by open space. Note the electric fields still go from one conductor to another, but now there are three conductors. (The complementary “elements” are the two wedge-shaped spaces, identical in shape to the conducting “elements” of the vee). The horizontal components of the electric field are anti-symmetrical around the bisecting normal plane, and thus cancel the horizontally polarized components in the forward or downward directions. The vertically-polarized components under the central electrode are only partially cancelled by those under the outer two electrodes, and thus radiation from this antenna is

vertically polarized (that is, the electric field is normal to the conducting interface in the far field). And the radiation is strongest in the direction defined by the intersection of the interface plane and the normal plane bisecting the vee angle. This radiation will then couple strongly to the dielectric slab waveguide mode if the high dielectric constant region is a slab rather than an infinite half space.

Note that the slot vee antenna requires a three-conductor feed. The source (or load) is connected between the conductor feeding the center element and the junction of the two outer elements.

Because these antennas are complements (or "duals") of one another, we may visualize the magnetic field lines of one antenna as being the same as the electric field lines of the other, and vice versa. Note also that while the sketches of the field lines in Fig. 5.1 are intended to illustrate the fields in the vicinity of the antenna elements, the fields around the parallel conductors feeding the antennas will have similar distributions. Thus the optical waveguides in the modulator would be placed symmetrically between the two conductors for the two-strip vee feed and the crystal oriented for a horizontal electric field (x-cut). Or the optical waveguides would be placed under the center conductor for the three-strip slot-vee feed and the crystal oriented for vertical electric field (z-cut).

The schematic design for Sheehy's concept is shown in Figs. 5.2-5.4, taken from the Rome Laboratories final report [Ref. 5.2]. Figure 5.2 shows a slot-vee receiving antenna coupled to a planar transmission line and terminated in a second slot-vee, which radiates and serves as a matched load to the transmission line. The "slots" or openings in the conductive plane are shown as black, with the white background representing the conducting plane (just the reverse of Fig. 5.1). The transmission line is "planar coax", with a center conductor and two ground planes. A portion of the center conductor is positioned over one optical waveguide of a Mach-Zehnder modulator. Figure 5.2 shows a series of such antenna and transmission line segments

20 GHz Slot-Vee Antenna Element

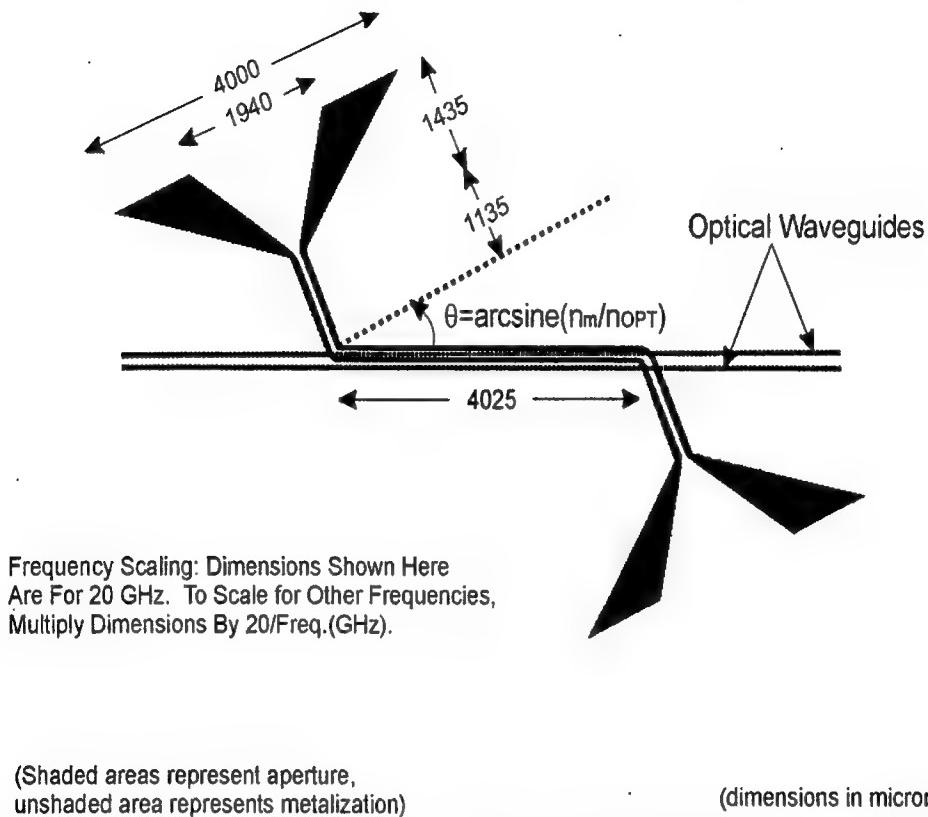


Figure 5.2 Schematic drawing of the slot-vee antenna elements overlaying one waveguide of a Mach-Zehnder modulator. (Redrawn from Ref. 5.2).

arrayed along the optical waveguides in the lithium niobate substrate. The mm-wave modulating signal is coupled in from the side from a slab dielectric waveguide into the lithium niobate crystal, also serving as a slab dielectric waveguide. (Or more specifically, it is an image slab dielectric waveguide, since one surface of the dielectric guide is completely covered by a

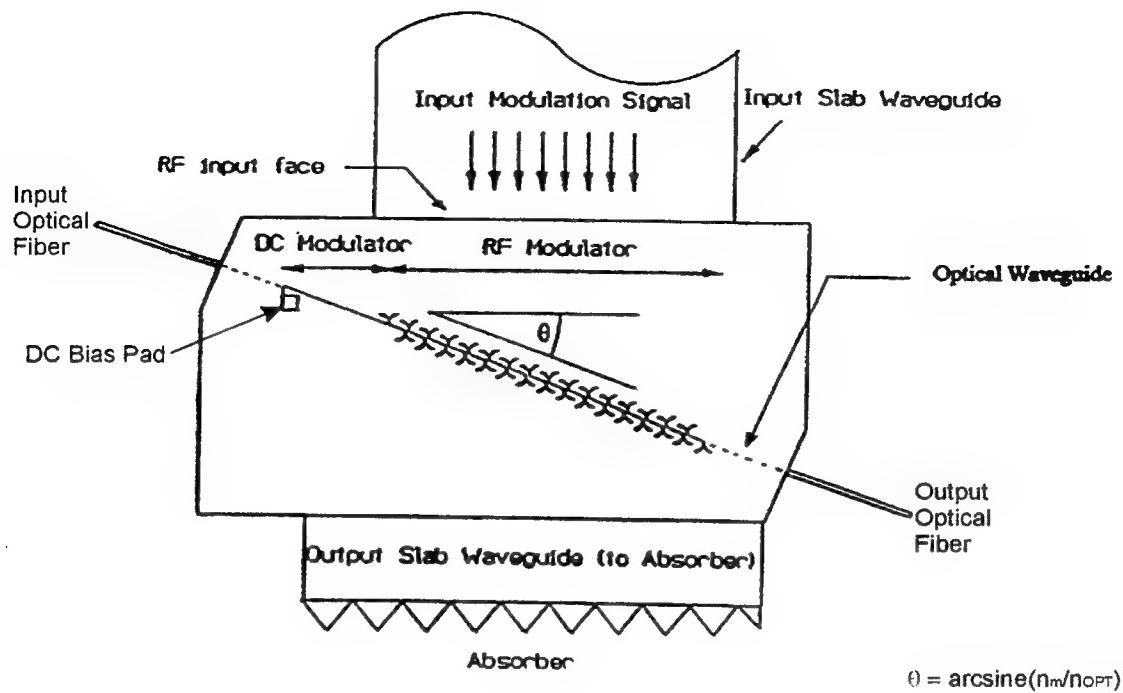


Figure 5.3 Schematic view of the slot-vee modulator. (From Ref. 5.2).

conducting layer, except for the slots of the antennas and transmission lines). One array of antennas receives the incoming wave, couples it along the transmission lines, then re-radiates it via a second array of antennas back into the slab waveguide, where it is absorbed at the far edge. Figure 5.4 shows Sheehy's concept of how the modulator could be mounted on a metal plane (conducting side down) and fed by a tapered image slab dielectric waveguide, fed in turn by a flared metallic half-waveguide.

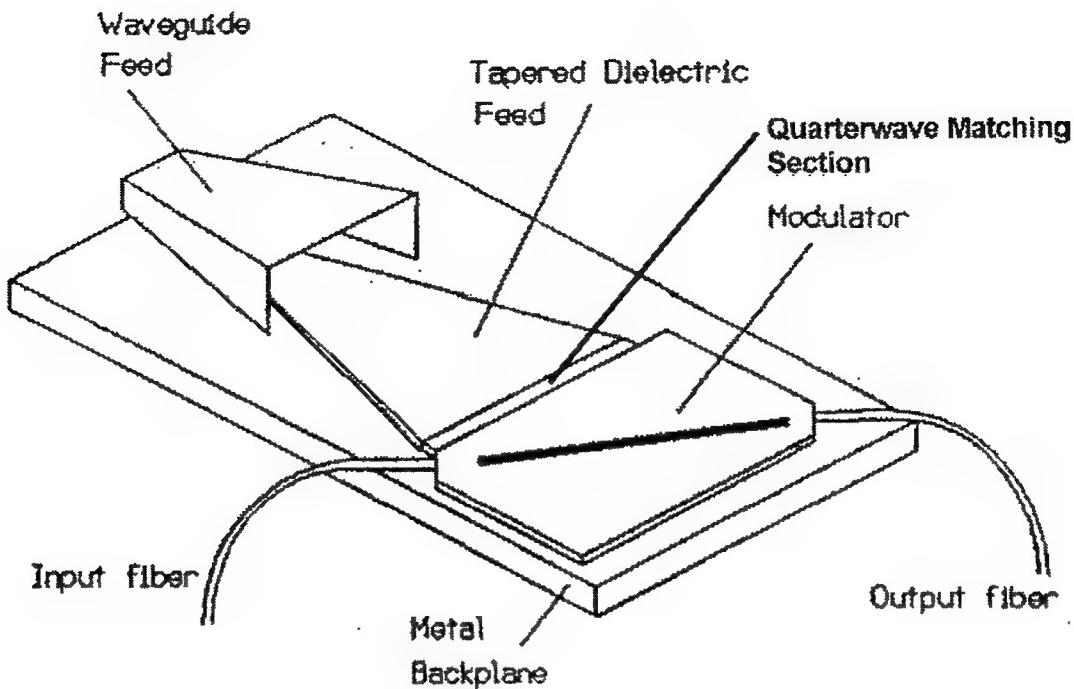


Figure 5.4 Schematic drawing of the slot-vee modulator and its microwave feed system. (From Ref. 5.2).

A U. S. Patent was also applied for and later issued [Ref. 5.4]. However, Sheehy was awarded his Ph. D. and left Caltech before executing a detailed design or any construction. Lee J. Burrows, a new graduate student that had joined the group in June 1992, was given the task in Fall 1993 of designing a slot-vee antenna-coupled modulator. The beginning of Burrows' work, as well as the latter portion of Sheehy's, was carried out with the support of a Rome Laboratory contract from November 1991 through November 1994, and is described in Ref. 5.2. During that contract it was decided to design waveguide and electrode masks that would yield modulators at 100 GHz, 44 GHz, and 20 GHz simultaneously. A copy of the electrode mask is shown in Fig. 5.5 (from Ref. 5.2). The masks were fabricated by Align-Rite and sent to the

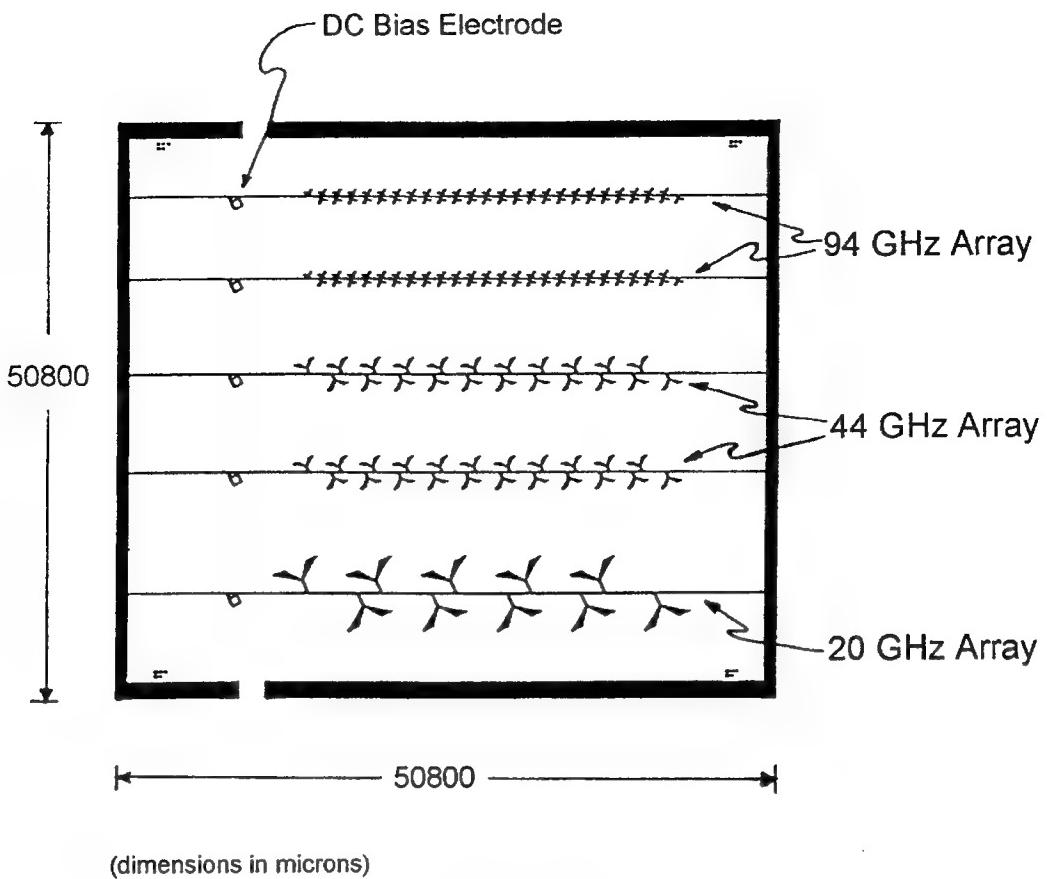


Figure 5.5 Antenna mask for 5 slot-vee modulators

Hughes Research Laboratories (HRL) for fabrication, under an informal arrangement with James H. Schaffner at HRL. Only one sample of this design was fabricated, but was received at Caltech after the contract had ended in November 1994.

5.2 Slot-Vee work during the funding gap period

Work on the slot-vee modulator was continued, but at a reduced pace during the gap in funding from November 1994 to February 1996, when the present contract began. During that gap time, some studies were undertaken at K-band (20 GHz) to develop the tapered dielectric image guide feed concept suggested by Sheehy (Fig. 5.4). Sheehy had envisioned a tapered image slab guide in a lower dielectric material (for example, teflon, with $\epsilon_r = 2$) expanding the fields to the desired width of the modulator, then coupling that wave to an image slab waveguide of lithium niobate, $\epsilon_r = 28$ (for z-oriented material), through a quarter-wave matching layer of an intermediate relative permittivity. This proved to be more difficult than we anticipated. Sheehy had used this arrangement successfully before in earlier wave-coupled modulators with full slab dielectric waveguides [Ref. 5.5]. But we found that the extra complication of making one side of the quarter-wave section precisely fit to the underlying ground plane of the image guide was very difficult indeed. And we also concluded that there was a significant amount of radiation

from the discontinuity introduced by the matching layer; i.e., it wasn't really "matched." This is one of the main difficulties in working with "open" waveguide structures.

In other measurements using WR-42 waveguide at 18-26.5 GHz, we attempted to make the tapered sections themselves of lithium niobate to avoid the fabrication and radiation difficulties of the quarter-wave layer. This proved impossible for another reason, namely that no matter what taper angle we tried, the wave was poorly coupled from an open ended metal waveguide into the tapered slab. This is just the opposite of what we observe with low dielectric constant tapers: Good coupling into a taper is generally very easy to obtain with low dielectric constant materials. However, these results did not come as a complete surprise, since much earlier work in this laboratory (unpublished, ca. 1979) showed similar difficulty in trying to couple from open-ended metal waveguides into tapered dielectric rods at X-band: With $\epsilon_r = 2$ to 10, it was trivially easy to obtain good coupling; with $\epsilon_r = 20$ to 32, it was impossible. No theoretical reason is known for this behavior, but it has now been observed in two very different geometries, so we conclude it is real. In any event, we abandoned the attempt.

During the funding gap we set up and measured the optical transmission of the HRL fabricated waveguides. The optical transmission was quite poor, about -20dB through the 2 inch length of the modulator with fiber butt-coupling (non-contacting) into the guide and lens coupling out to the detector. This optical loss was later found to be caused by "gaps" in the optical waveguides that resulted from short regions of the titanium evaporation "peeling off" before the in-diffusion took place. (Just the complement of the "mouse-bite" mask problem!) This phenomenon was observed in other modulators fabricated at HRL about the same time, so we were not alone in suffering. The problem was later solved at HRL by changing some processing steps.

Additionally, our waveguides were multimode. Evidently, the "prescription" for single-mode waveguides had changed when these guides were fabricated. It would have been nice to have had HRL fabricate another set of modulator chips using the same masks and the new process steps so that we would have had a modulator with low loss and single-mode operation.

However, during the funding gap, Jim Schaffner at HRL had become engaged in other technical assignments, and was not available to make additional modulator samples. We were thus forced to use the original modulator samples during the initial stages of the present contract, with unsuccessful results, as described later.

5.3 Initial Modulator Feed Design

It may turn out in the future that either or both of the microwave feed approaches studied during the “gap” period may be made to work with more experimental effort. However, shortly following the beginning of the present contract we decided that a better approach was to make the transition from air to lithium niobate *inside* the metallic waveguide, then expand it in a tapered lithium niobate image slab dielectric waveguide outside the metallic waveguide. Unfortunately, this required machine-shop facilities that were not affordable during the “gap” period, and had to wait until after the start of the present contract.

A modulator mounting structure was designed that had WR-42 input waveguide, then tapered down in height from 0.170 inches of the standard guide to 0.040 inches (1 mm) to match the thickness of the lithium niobate modulator chips. The width of 0.420 inches was held constant throughout the waveguide height taper. The taper design advice given by Moreno [Ref. 5.6] was “greater than two wavelengths.” We used about 4 guide wavelengths for the taper, which was 64 mm long. The taper was made by machining the top surface and side walls away, following a taper, then making a matching tapered cap that fit closely over the remaining WR-42 waveguide to form the tapered top surface. Both this cap and the cut down WR-42 waveguide then continued for another 25 mm at a constant 1mm height.

This tapered section was clamped in a machined channel in the top surface of mounting structure so that the lower inside plane of the tapered-down guide was continuous with the ground plane of the structure. This ground plane then extended to the area where the modulator would be mounted. A schematic drawing of the structure is shown in Fig. 5.6, and a photograph of the completed structure is shown in Fig. 5.7. Close-fitting plugs of low-loss dielectric material were machined to fit in the 0.040 x 0.420 inch cross section of the guide. These plugs could be placed in the guide by removing the top cap of the guide.

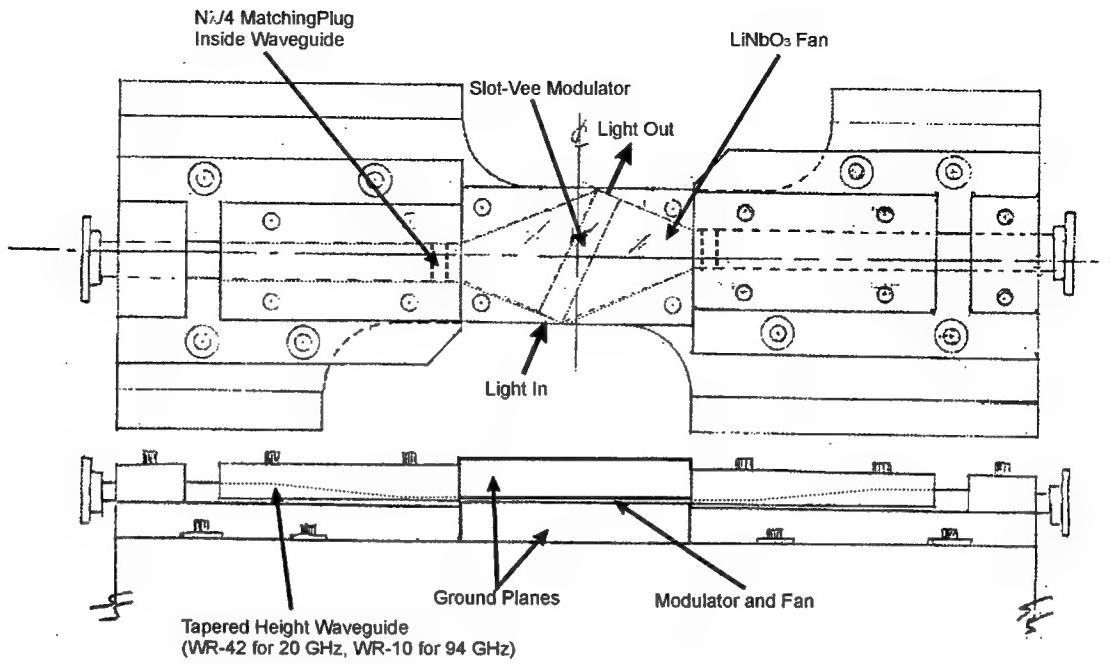


Figure 5.6 Assembly drawing of modulator support and feed structure.

Two different sets of plugs were initially fabricated:

Set 1. A single matching section of $\epsilon_r = 4$ Stycast® artificial dielectric, 0.237 inches long (which is three quarters of a guide wavelength at 20 GHz in the reduced height guide.)

Set 2. A pair of matching sections of $\epsilon_r = 3$ and $\epsilon_r = 10$ Stycast®, 0.279 and 0.143 inches long respectively (both three quarters of the guide wavelengths at 20 GHz in the reduced height guide.) Neither of these sets make perfect matching transformers, but represented the best we could do with stock values of artificial dielectric constant material.

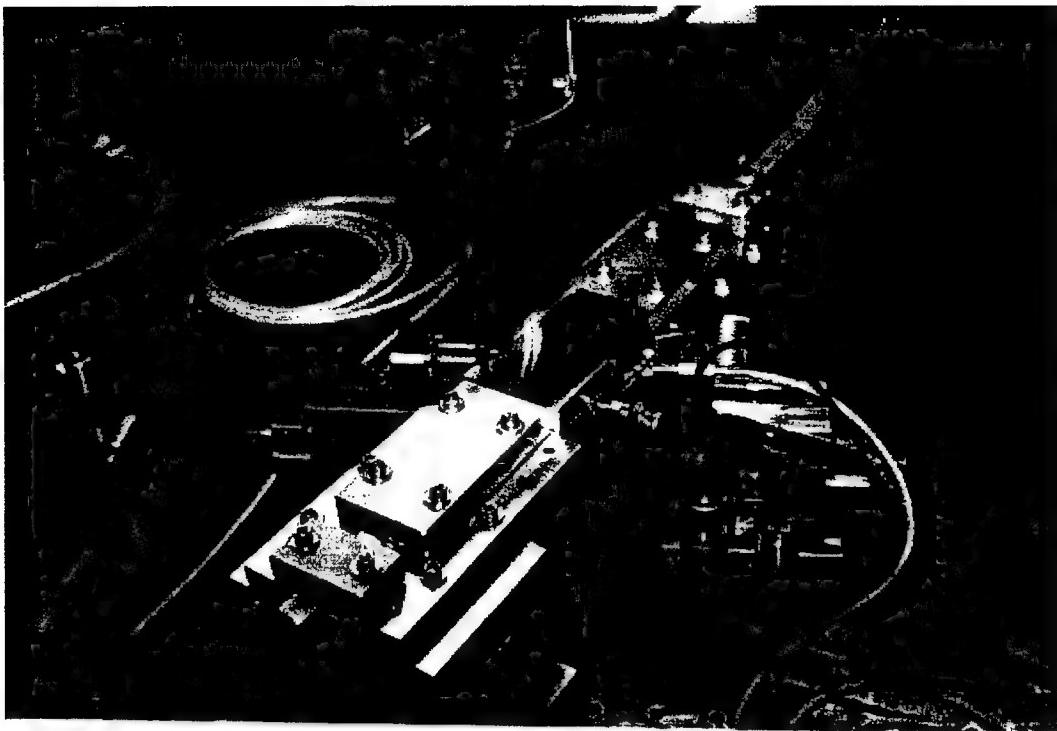


Figure 5.7 Photograph of the slot-vee modulator test setup, with fiber input coupling (from the left), and lens output coupling (to the right).

In order to check the effectiveness of these tapered sections in matching to lithium niobate, and to determine the overall transmission of the whole feed design, the modulator mount was made in a symmetrical fashion, with an output coupling waveguide made in the mirror image of the input guide. In this way, we could measure the microwave transmission through the entire system, and then argue that the loss from the input to the modulator would be about half (in dB) the loss through the entire system. The distance between input and output metal waveguides could be changed by loosening the clamps, so that the transmission could be measured for various combinations of elements in the path.

The results of an early single-frequency measurement are given in Table 5.1. The values listed in dB are the values measured flange-to-flange between the input and output WR-42 waveguides with the components listed inserted in the mount.

We can then argue that the transmission from the input flange to the modulator is about half the total, or about -8 dB. However, we need some caution here. We can interpret the results given in the table as follows: The empty waveguide loss in the tapers is likely due to (1) losses caused

by the lack of fit in the removable top cap of the tapered section and (2) losses butting the two reduced height ends together without flanges.

Table 5.1
Measured Flange-to-Flange Loss in the Modulator mount

Empty tapered waveguides:	-3.6 dB
Tapered waveguides with $\epsilon_r = 4$ plugs inserted	
flush in the ends	-5.8 dB
As above, with lithium niobate plugs also	-5.8 dB
As above, with lithium niobate fans added	-15.5 dB
As above, with modulator chip added between the fans	-16.1 dB

The loss tangent of Stycast® material at 20 GHz is less than 0.001, so there should be negligible loss in a length the order of a wavelength. There should be no surface reflections since the two plugs together are and integral multiple of a half guide wavelength. The 2 dB additional loss with the plugs can only be attributed to reflections from the small airgap where they contact each other.

Adding the lithium niobate sections did not produce additional loss, indicating that the $\epsilon_r = 4$ plugs did a good job of matching in and out

The fan-out and fan-back-in resulted in about 10 dB additional loss. Adding the modulator between the fans produced little additional loss. This 10 dB loss from the fans is a serious amount. However, it was not unexpected. The input fan very likely produces a somewhat cylindrical wavefront, just as a flared horn antenna does. It is only in the far field of the antenna that the curvature decreases and the wave becomes essentially plane. But here, the modulator and the output fan are not in the far field. The input to the output fan would like to see the mirror-image curvature of the phase fronts for good transmission. Unfortunately, it sees the opposite curvature, so that a good wavefront match is not obtained. This will result in reflection and radiation from the fan.

We could argue that the output fan is not used in the actual modulator, so that the transmission loss from input to the modulator should not suffer one half the 10 dB due to the fans, but something substantially less. However, the antenna array on the modulator will also be affected by the curvature from the input fan, so a conservative estimate is that -8dB transmission from input to modulator is probably a good one.

The way to reduce the loss (to something more like -3dB) would be to correct the phase front from the input fan with a "lens" or a better flare design. A "lens" could be effected by adding a shaped piece of material onto the top of the waveguide to change the guide velocity slightly. However, this may also introduce some additional reflection and radiation from the guide. We elected to live with the -8 dB estimated loss for our experiments and leave the additional improvements for later.

Further measurements on the flange-to-flange transmission were made using a Hewlett-Packard hp 8722D network analyzer over the range 18-26 GHz. These measurements showed that there was considerable (more than 10 dB) variation in the transmission with frequency. The rapidity of the variation with frequency suggested that the cause was interference between sources of reflection that were separated several cm in space. For example, the various butt joints between the artificial dielectric and the lithium niobate became suspect. These measurements also showed a variation in the fine structure with applied pressure to the combination of dielectric pieces in the assembly, also indicating that the small gaps between sections were the source of the interference.

5.4 Measurements with the Original HRL Modulator Chip

The first slot-vee modulator substrate was received from HRL during the contract gap period, with the high optical loss and multi-mode behavior as mentioned above. In addition, some modulator samples had shorted bias transmission lines. We determined that the cause of the shorts was a thin residual layer of chrome plating between the bias line and the surrounding ground plane. (Similar problems were discussed in Section 3 with the DCM.) We suspect that the central r-f transmission line segments may also have been shorted for the same reason, but we could not measure this for a fact, since that line is connected to the ground plane through the slot-vee antenna. In any event, we elected to etch off all the metal coatings and re-plate the electrodes ourselves. This was done successfully to obtain a chip in which the bias line was

short-free (and we assume the r-f lines were also, although we could not confirm this by dc measurement).

The result of d-c measurements on the bias line gave a V_{π} of 81 volts for a 6 mm bias line length. The extinction ratio (maximum transmission divided by minimum transmission, in photocurrent) was less than 3 dB as a result of the multimode nature of the optical waveguides. All in all, this was a very poor chip, but it was all we had at the time.

5.5 First Modulation Tests

About a year into the program, we attempted to obtain modulation with the original HRL fabricated chip and the feed structure described in Sec. 5.3 above. We used a Hewlett-Packard hp 694 BWO sweeper with +10 dBm output over the range 18 to 25 GHz . We attempted to see modulation sidebands using a scanning Fabry-Perot interferometer (SFP), and also with a New Focus model 1011 photodetector, which has a 45 GHz roll-off frequency. A Hewlett-Packard hp 8559A spectrum analyzer was connected to the detector output. We were unsuccessful in seeing modulation with either detection method. We considered several possible causes:

- (1) Poor optical transmission (worse than 20 dB optical loss through the modulator);
- (2) Poor slope of the modulator transmission curve because of the poor extinction ratio from the multimode waveguides.
- (3) Poor coupling from the slab waveguide to the slot vee antennas;
- (4) Insufficient detection sensitivity.

At the time, we did not consider the possibility that the region between the transmission line electrodes was still shorted or lossy because of residual un-etched metal, but in light of the lack of modulation seen in the DCM described in Sec. 3, that also now seems a possibility.

Items (1) and (2) could only be changed with a new modulator chip, which we were attempting to obtain from HRL. Item (3) was studied with a side experiment described in the next section. Item (4) was addressed by obtaining an r-f signal source with 15 dB more output power.

5.6 Antenna Coupling Experiments

We decided to investigate the coupling of a wave in a dielectric image line to a slot-vee antenna experimentally. We are unaware of any reasonable quantitative theory for this antenna, so an analytic approach was out of the question. Mousessian's study had been experimental, and carried out on relatively low dielectric constant substrates, (fused silica, $\epsilon_r = 4$) while we were using z-cut lithium niobate with $\epsilon_r = 28$. A further worry in our experiment was the physical thickness of the substrate was only 1mm, while the distance between the ends of the vee slots was 3.8 mm for the 20 GHz antennas. Thus the approximation of an "infinite half-space" for the 1 mm thick substrate was not a good one for this sample. (The corresponding dimension for the 94 GHz antennas was 0.7 mm, so the approximation would seem to be much better). We did not know if this "cramping" of the fields surrounding the antenna elements would result in poorer coupling to the guided wave in the dielectric or not, so we attempted to determine the behavior experimentally.

We decided to do the experiment with a 10-times larger model at 2 GHz. We cut two slot-vee antennas into the corners of a foot square piece of copper foil, as shown in Figs. 5.8 and 5.9. The antennas were fed by short pieces of teflon-insulated rigid coax soldered to the foil. The foil was placed atop a foot square pieces of Emerson and Cuming Stycast® $\epsilon_r = 30$ artificial dielectric. The thickness of the dielectric could be varied in increments by stacking several pieces of the Stycast® material. The $\epsilon_r = 30$ slabs in turn were placed atop a piece of polyfoam a foot square, or a piece of high dielectric constant absorbing material, Emerson and Cuming type HF, which has a loss tangent of order unity, and a dielectric constant of about 26 in this frequency range, or a second sheet of copper conductor. Sitting on the polyfoam, the region between the antennas looks like a slab dielectric waveguide. Sitting on the type HF material, the region looks like an infinite half-space of $\epsilon_r = 30$ material. Sitting on the conductor, the region looks like a parallel-plate metal waveguide filled with $\epsilon_r = 30$ material. The transmission from coax-to-coax was measured on the hp-8722D network analyzer.

An estimate of the transmission loss that can be expected from two antennas connected by a dielectric slab or parallel-plate metal waveguide can be obtained from the two-dimensional analog of the famous Friis formula for path loss in free space between two antennas. To find this formula, we make the following assumptions:

- (1) The antennas couple perfectly to the slab.

- (2) The wave from an antenna expands cylindrically, decreasing as $1/R$ in power (rather than $1/R^2$ for a spherical wave in free space).
- (3) The antennas have unity gain. (The actual gain of our slot vee antennas should be greater than this, since their arms are about 0.6 wavelengths long at 2.4 GHz).
- (4) The antennas are in each other's far field.

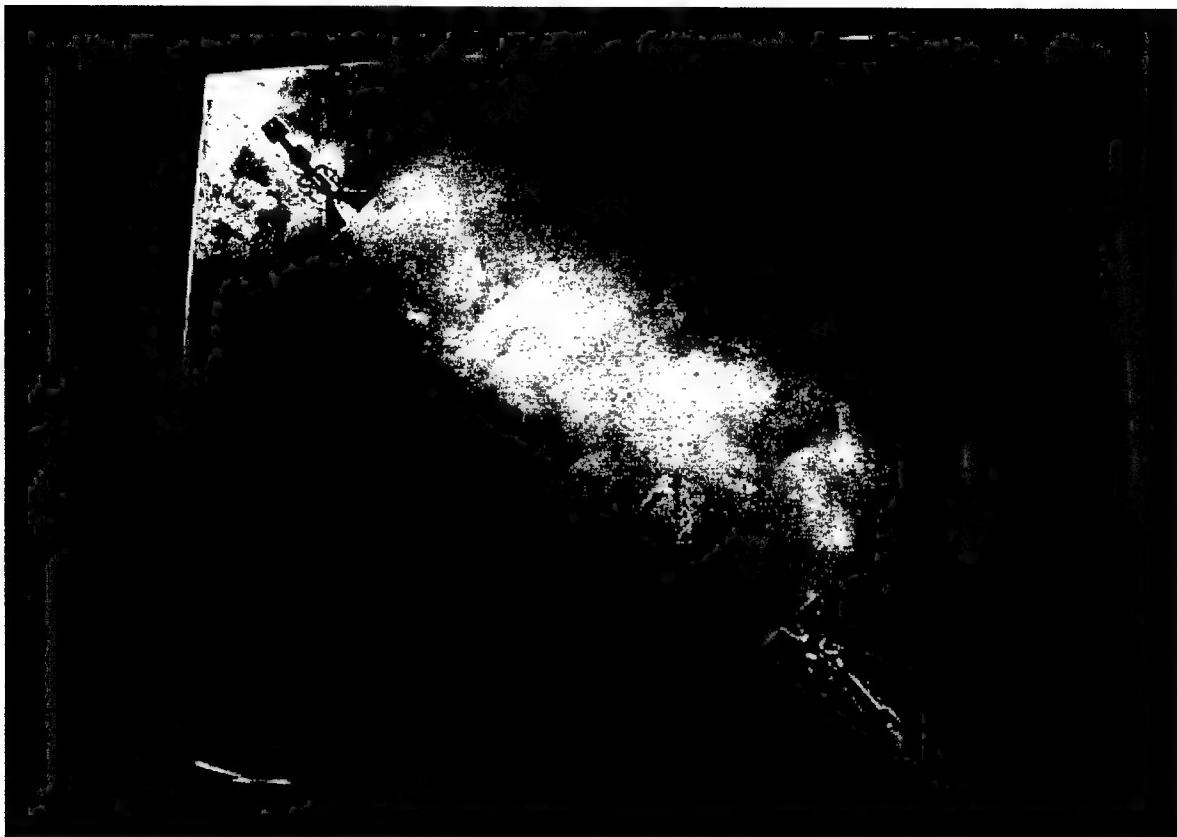


Figure 5.8 Overall view of the scale model transmission setup, showing two slot vee antennas.

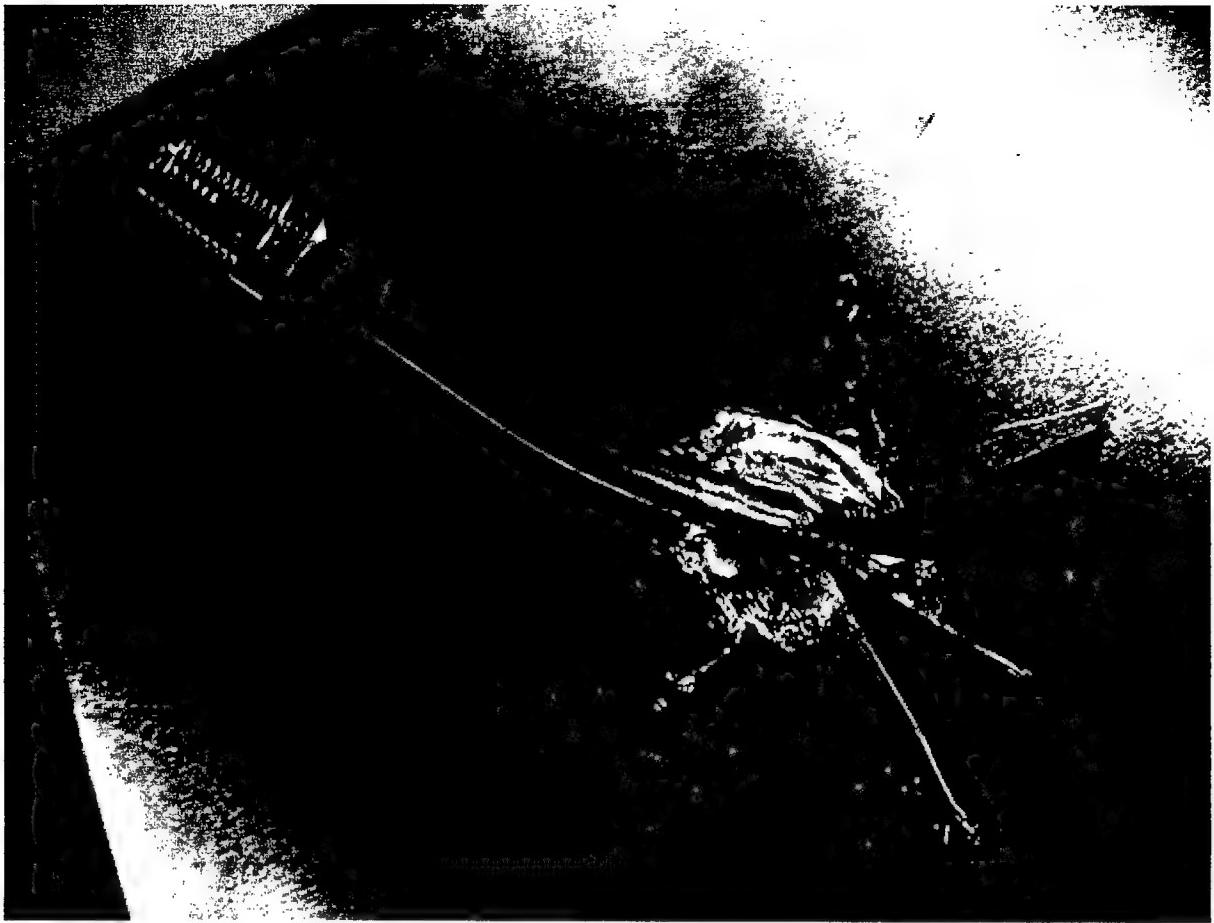
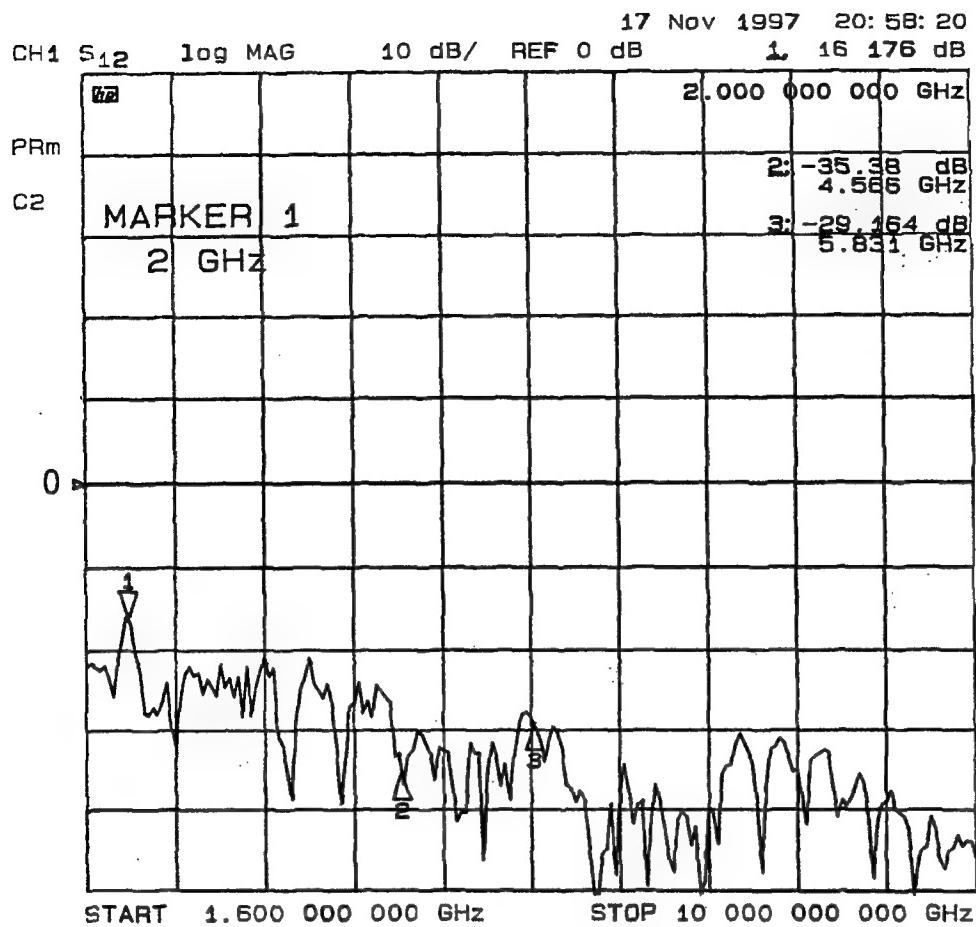


Figure 5.9 Close up detail of one slot vee antenna with coax line feed.

With these assumptions the two-dimensional Friis formula gives a transmission power loss of about $\lambda / (2 \pi)^3 R$, where λ is the wavelength in the dielectric medium and R is the separation between the antennas. For a frequency of 2.4 GHz in $\epsilon_r = 30$ material, the wavelength is 22.8 mm. In our setup, the antennas were spaced 240 mm. Thus, we would expect a path loss of about 34 dB between isotropic antennas. Our slot-vee antennas actually have some gain, which will reduce the path loss.

We measured approximately 20 dB path loss as shown in Fig 5.10, for a 6.7 mm thick substrate (our thinnest) over polyfoam. This corresponds to a lithium niobate slab thickness of 0.67mm, smaller than the 1mm substrates actually used.



**Figure 5.10 Transmission between slot vee antennas on a 6.7 mm Styrofoam®
 $\epsilon_r = 30$ substrate**

Figure 5.11 shows the same measurement with a 28.7 mm thick stack of $\epsilon_r = 30$ plates, the thickest we could produce. We note that the two thickness give almost the same resulting transmission, -20 dB. This comparison tells us that the antennas are working properly, and their fields are not "cramped" by a thin slab waveguide, even a slab waveguide that is substantially thinner than the spacing between the outer ends of the antennas (27 mm here).

The 34 dB – 20 dB = 14 dB difference between theory and experiment would require that the antennas have 7 dB gain each. This seems a little high for a vee antenna that is only 0.75 wavelengths on a leg (assuming the effective wavelength is 22.8mm) or 0.6 wavelengths on a leg (assuming the effective wavelength is given by the average of the air and dielectric media,

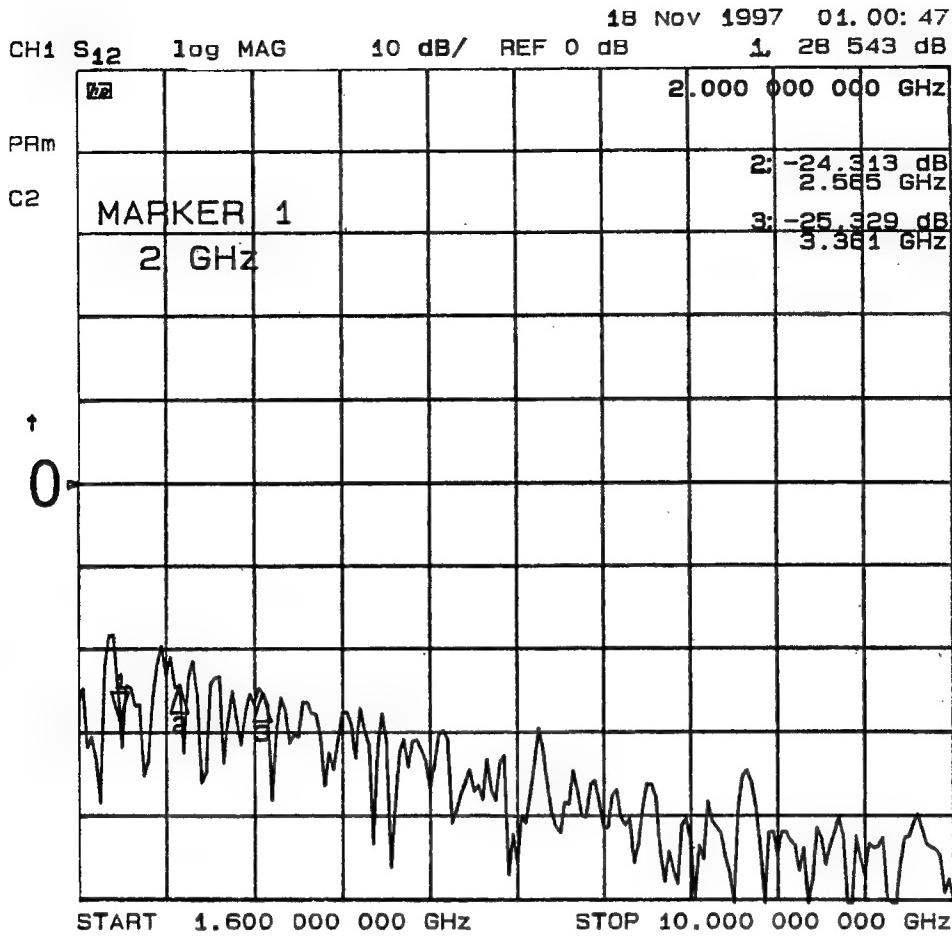


Figure 5.11 Same as 5.10, but 28.7 mm thick

(32 mm). Something of the order 3 or 4 dB seems more reasonable for the gain of such a vee antenna. However, since we have no theory for the wedge-element slot vee, we cannot make any definitive conclusion about what its gain should actually be.

Similarly, varying the material on the far side of the dielectric slab from air to lossy dielectric to metal had little effect on the average transmission, only the details of the fine structure. We conclude that slot-vee antennas couple very well to a dielectric slab, even if that slab is thinner than the separation between the arms of the vee, thus “cramping” the EM fields around the antenna.

5.7 Modified Coupling Components

Laboratory signal generators typically have the same output power as our hp 694, so borrowing a different generator from another group would not help our measurement. And high power K-band amplifiers are not that common. However, reflex klystrons with powers of + 15 dBm or more are not uncommon. We happened on an advertisement for "new old-stock" surplus Varian VA-98M (JAN-8901) reflex klystrons from Fair Radio Sales of Lima, Ohio. These tubes have a limited tuning range (23.6 to 24.4 GHz) and a power output in excess of 300 mW (+ 24 dBm). They were \$200 each. This seemed like the cheapest way to get more power at K band. And we had the power supply to run them.

We ordered two tubes. They were still sealed in their original cartons, with manufacture date of October 1980. Unfortunately, both tubes proved faulty on testing. They had become gassy in the past 17 years of storage. We returned the two tubes for credit and ordered two more. These also were from the same manufacturer's lot. One exhibited a "hot" cathode-to-body short (i.e., it was open when cold, but shorted as the cathode expanded on heating). However, the other tube tested good, and produced 300 mW at 24.0 GHz.

We re-measured the transmission of the waveguide matching structure over the range 18 to 26 GHz using the hp-8722D. The end-to-end loss of the two empty tapered waveguides alone varied from 1.1 to 2.0 dB over this range. Unfortunately the losses through the matching sections with the $\epsilon_r=4$ three-quarter wave plugs were about 3 to 4 dB worse at 24 GHz than at 20 GHz, as seen in Fig 5.12.

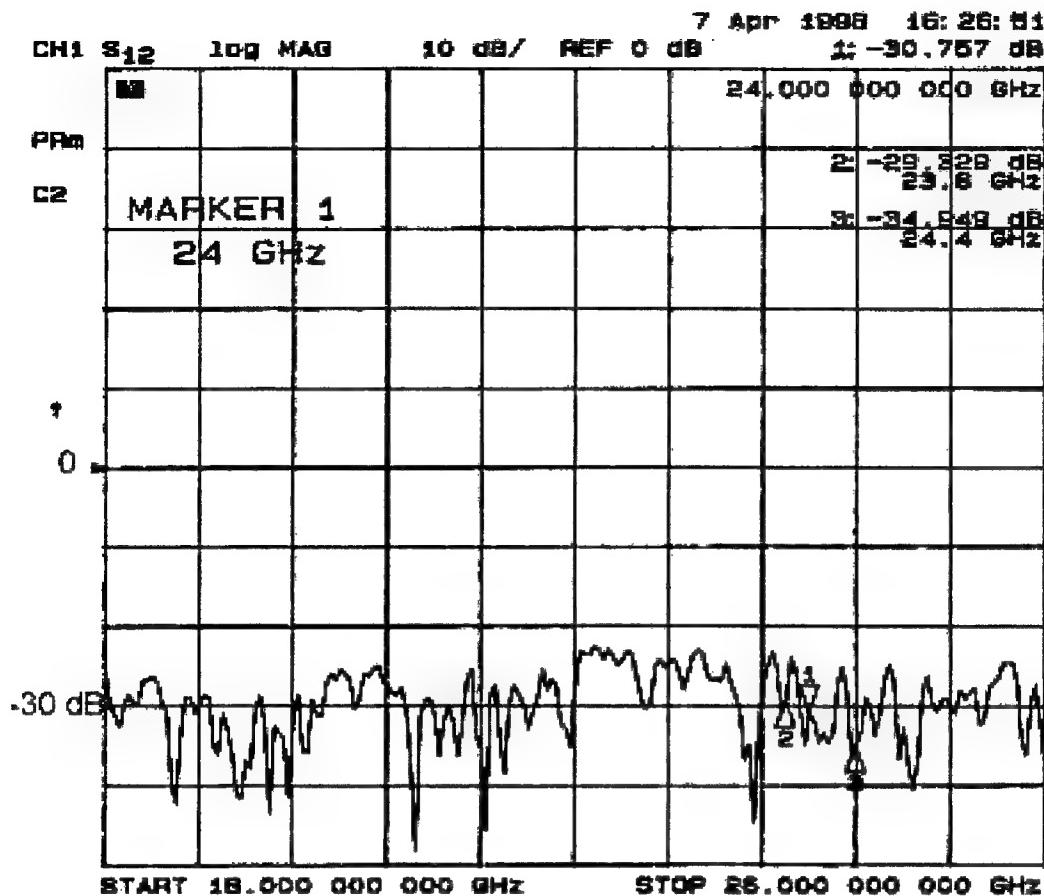


Figure 5.12 Transmission 18-26 GHz with old 20 GHz plug set.

We decided to have new matching plugs fabricated to try to reduce this loss. Two sets of plugs were made: single three-quarter wave plugs of $\epsilon_r = 5$ material and dual three-quarter wave plugs of $\epsilon_r = 2.5$ and $\epsilon_r = 11$ material. The two-plug set gave the better result, -23 dB transmission at 24 GHz for the entire assembly of plugs, lithium niobate fans and modulator chip. This was 7 dB better than the transmission at 24 GHz using the original 20 GHz plugs.

During the measurements, we tried an additional experiment, namely, placing a metallic plate over the top of the assembly of plugs, fans, and modulator. The transmission system thus became a dielectric-loaded parallel-plate waveguide rather than an open dielectric image guide. Originally, this approach was rejected, since in theory the metal plate would produce additional loss due to its finite conductivity, while the dielectric-air interface acts as a lossless reflector. However, in practice, all of the butt joints radiate and contribute loss. The metal plate eliminates these radiation losses (but may simply turn them into reflections). In this configuration, the

added metallic plane seemed to help (just the opposite of the results obtained in Sec. 3). We had a precision plane fabricated to fit the open guide region. With this plane in place the losses at 24 GHz decreased from -23 dB to -17 dB. Furthermore, much of the "fine structure" disappeared, giving a much smoother transmission curve with frequency as seen in Fig. 5.13.

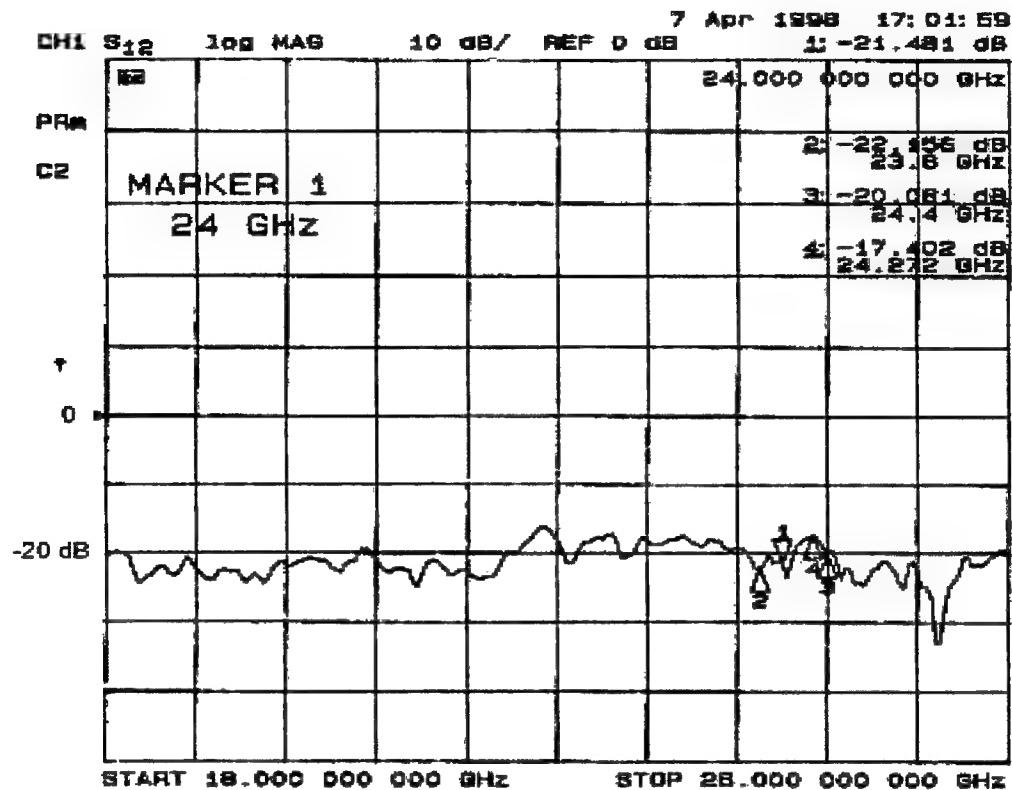


Figure 5.13 Transmission curve with 24 GHz plugs and top plane.

5.8 New Modulator Chips

About a year into the present program, HRL began fabrication of some new titanium in-diffused optical waveguides for us. We agreed that they would make only the optical waveguides, and we would put the silicon dioxide layers and metal electrodes on at Caltech. We also decided to make individual modulators rather than the entire mask shown in Fig. 5.5, since we were concentrating on getting the 20 GHz modulator to work and didn't need the additional 44 and 94 GHz modulators.

The first 0.4 inch by 2 inch substrates with in-diffused lithium niobate waveguides were received from HRL in June 1997. We had to modify the holding fixture on our photoresist spinner to

accommodate these rectangular samples (so they don't fly off while spinning!) SiO₂ buffer layers were sputtered on using the reactive sputtering system in Prof. Axel Scherer's laboratory. Uniform layers of chromium (200 Angstroms) and gold (1200 Angstroms) were sputtered on top of the SiO₂ layer, and photo resist on top of this. After exposure and developing of the photoresist, the gold and chromium not covered by resist was removed by reactive ion etching (RIE) in an argon plasma. Finally, the remaining photoresist was removed.

Modulator chips made by this method were evaluated with dc bias voltage only applied to the bias lead. We noted substantial bias drift, more than 20% of V_π in a few seconds, with voltages of about 20 V applied. This severe drift was "cured" by annealing at 400 C for 60 minutes, with a 1 C/min ramp up and down.

After annealing the drift was much reduced (but not reduced to zero). We measured a V_π of 20 V in the 6 mm long dc bias section (contrasted with 81 V for our old modulator chip!). The maximum optical transmission was -8 dB (much better than the original modulator samples of 1994) and biased at minimum was -11 dB, for an extinction ratio of about 3 dB, about the same as the earlier modulator. The waveguides were obviously still multi-mode, and we conclude that too much titanium had been used in making the guides at HRL. However, there was no chance of obtaining still another run of waveguides from HRL, so we just continued the best we could.

5.9 Detection Sensitivity Improvements

Our primary means of detection for modulation is the scanning Fabry-Perot interferometer display of the modulation side bands. While we can also use the New Focus fast detector at 24 GHz, it doesn't reach 94 GHz. And it has also has a severe limitation in the average photo current before burnout. Since we have much more laser power available (175 mW), we can do better with the Fabry-Perot and a much more rugged (but slower) photodetector.

The limitation on the scanning Fabry-Perot sensitivity is the noise level displayed on the oscilloscope. The modulation sidebands must significantly exceed this noise level to be seen. We had been using the built-in photodetector amplifier in the Spectra-Physics 470 scanning unit, but this proved to be substantially noisier than simple shot noise from the photodetector. Thus there appeared to be some room for improvement by building our own amplifier, as we previously described in Sec. 3.4.

The noise performance of operational amplifiers (op-amps) is usually characterized by an equivalent noise voltage source in series with the input and an equivalent noise current source in parallel with the input. These two sources are assumed to be independent. Thus, if the op-amp input is open circuited, the resulting noise comes entirely from the current source, and the voltage source contributes nothing. Conversely, if the op-amp input is short circuited, the resulting noise comes entirely from the voltage source, and the current source contributes nothing. If the input is terminated in a finite resistance, then both sources contribute, and their relative contributions are determined by the value of the resistance termination.

Since our photodetector behaves as an ideal current source and thus essentially an open circuit as far as the input of the op-amp is concerned, it is important to have an op-amp with low values of equivalent current noise, while high values of equivalent voltage noise do not contribute. This is the opposite of what is usually meant by a “low-noise op-amp”, where the input is driven by some low impedance source. Thus, we searched for a readily obtainable low current noise op-amp, and selected the Analog Devices 795 op amp. We built a two stage amplifier, using this op-amp as the input stage, and a conventional “low-noise op-amp,” the Analog Devices 797, as the output stage, since it is driven by the low impedance output of the input op-amp. This amplifier proved superior to the unit built into the S-P 470. It improved our detection sensitivity considerably.

5.10 Final Experiments

Armed with the improvements detailed above, we attempted to obtain modulation at 24 GHz. An optical input power of about 20 mW was used for these experiments. The klystron was tuned over the range 23.8 to 24.4 GHz. The frequency was measured accurately by heterodyning a sample of the klystron output with the second harmonic of an X-band signal generator, whose frequency was accurately measured with a wave meter (hp-537). The output power of the klystron ranged from 150 mW at the low frequency end to 310 mW at the high frequency end (as measured with a FXR bolometer detector through a 10 dB directional coupler and 10 dB attenuator).

We observed modulation sidebands, as shown in Figs 5.14 and 5.15. The modulation sidebands are the extremely small “pips” in the (a) parts of each figure. The multiple signals in the (b) parts are the multiple modes of the ATX laser, a nuisance in the measurements, since they obscure a portion of the modulation tuning range.

Because the scanning Fabry-Perot has a free spectral range of 2.0 GHz, we cannot see side bands near even multiples of 2.0 GHz, e.g. 24.000 GHz, since these coincide with the carrier aliases. But we can see the sidebands away from this frequency if they are strong enough to rise above the "skirt" of the carrier aliases. As a "sanity check", the pips were observed to move in frequency correctly as the klystron frequency was tuned over its oscillation range.

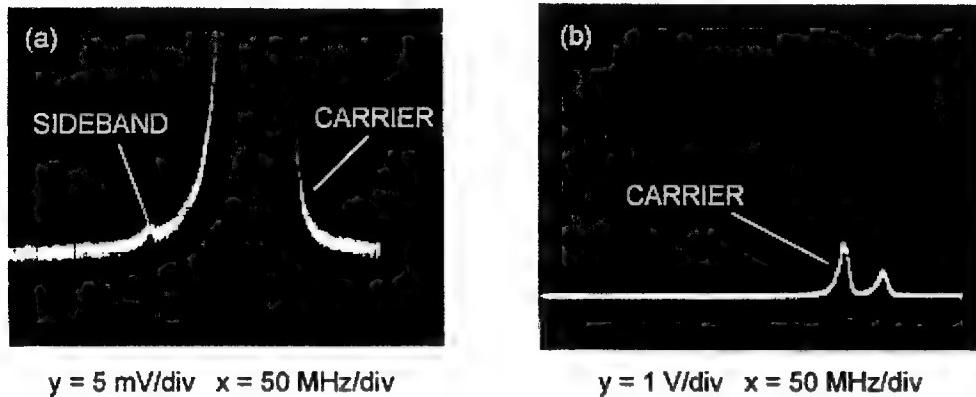


Figure 5.14 (a) modulation sideband at 23.472 GHz and carrier (off scale)
(b) Carrier on scale. Note different oscilloscope sensitivity.

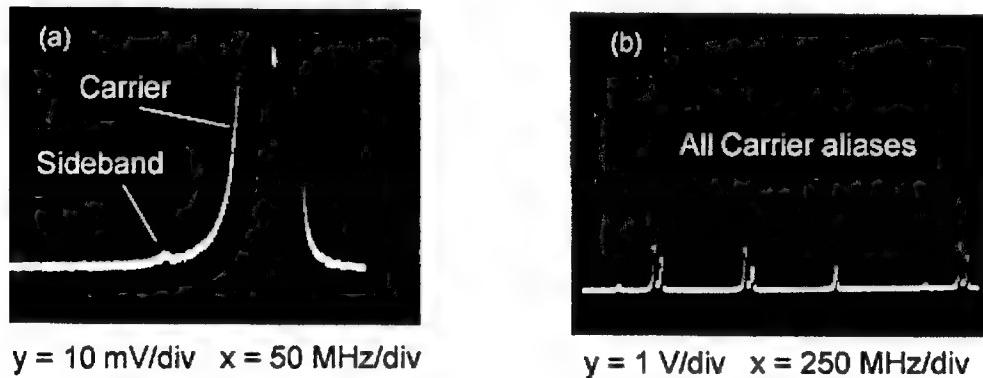


Figure 5.15 (a) Modulation sideband at 23.472 GHz (carrier off scale).
(b) Carrier on scale. (Note different oscilloscope sensitivity).

By taking the ratio of the sideband amplitude to the carrier amplitude, we may determine the percentage modulation from a simple formula [see Ref. 5.1].

$$m^2 = 16 P_{\text{SIDEBAND}}/P_{\text{CARRIER}}$$

where the powers are the optical powers and thus photocurrents as seen on the scanning Fabry-Perot display.

The figure of merit for the modulator is the ratio m^2/Watt . For the range of sideband amplitudes measured this value ranged from 0.09 to 0.17 [m^2/Watt].

5.11 Discussion of the K-Band Modulator Results

As a point of comparison, a standard electro-optic modulator characterized by its value of V_π and transmission line impedance R_m has a modulation index m when driven with a r-f voltage V_m given by

$$m = (\pi V_m / V_\pi)$$

and requires a r-f drive power P_m to do so, where

$$P_m = (V_m^2 / 2 R_m)$$

Thus the same figure of merit for a conventional modulator is

$$m^2/P_m = 2 \pi^2 R_m / V_\pi^2$$

Thus an ideal lossless conventional modulator with $V_\pi = 10 \text{ V}$ and $R_m = 50 \text{ Ohms}$ would have a figure of merit of 9.8 [m^2/Watt], about 54x better than our observed result. However, the result compares reasonably with Sheehy's 94 GHz bow-tie modulator result of 0.072 [m^2/Watt] [Ref. 5.5] and Cummings' result of 0.2 [m^2/Watt] for the 94 GHz directional coupler modulator [Ref. 5.7].

Recall that we measured 23 dB r-f loss flange to flange in the modulator mount. If we assume that half this loss exists between the input flange and the modulator antenna array, then we can assign 11.5 dB loss, or a factor of 14 to the input coupling. If we back out this factor, then the present modulator is only 3.8x worse than our comparison 10 Volt conventional modulator. In addition, two of the five antenna / transmission line sections on the modulator were shorted and thus likely totally ineffective. Our modulator should have been $5/3 = 1.7$ times higher in performance had all antennas been effective. Thus we are 2.3x worse than our comparison 10 V modulator.

A further correction of unknown magnitude should be made for the multimode operation of the optical waveguides. The phase modulation of the higher order modes in the arms of the interferometer does not result in amplitude modulation of the output. But the feed-through of these modes does contribute to $P_{CARRIER}$ in the formula above. Thus the experimental value of the ratio of sideband to carrier power is reduced by these higher order modes. Operation with single mode waveguides should yield a higher value of m^2/Watt , even with all the other defects mentioned above present. It is difficult to estimate the degradation from multimode waveguides. It would be best to re-fabricate the modulator chips with single mode waveguides, and short-free antennas.

5.12 Modulators at 44 and 94 GHz

At the outset of this program, we intended to use the K-band modulator as a "large scale model" to work out the coupling and mounting schemes for the modulator chip. However, unless a very long modulator is needed, a conventional electrode structure is likely more practical at this frequency, since the effects of velocity mismatch are only beginning to be felt at 24GHz and 10 to 20 mm modulator lengths. The main interest was in applying the proposed slot-vee configuration to W-band modulators, where the penalty for velocity mismatch in this length modulator is much more severe. For that reason, the mask design of Fig. 5.4 included 44 GHz and 94 GHz modulator electrodes.

Tapered height waveguides were fabricated in WR-10 waveguide, and an initial set of odd quarter-wave matching plugs were fabricated, as was a set of lithium niobate fans to expand the excitation from the 0.100 inch width of the waveguide to the 2 inch width of the modulator. Some preliminary transmission measurements were made with these waveguides mounted in the modulator support structure (during a slack period awaiting delivery of K-band modulator chips). The losses measured were somewhat worse than the 23 dB at K-band, but not a lot worse. We did not have much time to optimize the coupling scheme. These were single-frequency measurements with a W-band reflex klystron, since the HP-8722D does not operate in W-band. Unfortunately, one of the lithium niobate expansion fans was broken during the measurements, and we were unable to have another one cut on the dicing saw at HRL. (The fans are much more time-consuming to cut, since they require a saw set-up at odd angles).

The only modulator chips completed at 94 GHz were the original batch, which had defective optical waveguides (gaps from the titanium “peel” and multimode operation). We ran out of time on the contract before fabricating new 94 GHz guides.

Given the problems encountered at 24 GHz, we were not particularly enthusiastic about pursuing the 94 GHz modulator. This was also true of the 44 GHz modulators, for which we had only the original defective modulator chips.

5.13 Conclusions

We have successfully demonstrated the slot-vee antenna-coupled modulator originally proposed by Sheehy (Refs. 5.1, 5.4) at 24 GHz. While the experimental performance obtained was substantially worse than commercial conventional 24 GHz modulators available today, there were enough known defects in the fabrication of the modulator chip and the RF input coupling scheme to estimate that comparable performance would be obtained upon correction of the defects. The results of our studies were presented at the SPIE meeting in San Diego in July 1998 [Ref. 5.8].

An RF feed structure was built for 94 GHz operation and partially evaluated. Modulator chips for 44 GHz and 94 GHz were fabricated but were found to have serious defects. The program ended before additional 94 GHz chips could be fabricated.

5.14 References for the Slot-Vee Antenna-Coupled Modulator

- 5.1 F. T. Sheehy, “Antenna-Coupled mm-Wave Electro-Optic Modulators and Linearized Electro-Optic Modulators,” Ph. D. dissertation, California Institute of Technology, June 1963.
- 5.2 W. B. Bridges, L. J. Burrows, U. V. Cummings, R. E. Johnson, and F. T. Sheehy, “60 and 94 GHz Wave-Coupled Electro Optic Modulators,” Final Report RL-TR-96-188 on contract F30602-92-C-0005, with USAF Rome Laboratory, September 1996.
- 5.3 A. Moussessian and D. B. Rutledge, “A Millimeter-Wave Slot-V Antenna,” Proc. IEEE AP-S International Symposium, Chicago, Illinois, 18-25 July 1992, Vol. 4, p.p.1894- 1897.
- 5.4 F. T. Sheehy, U. S. Patent No. 5,309,531, issued May 3, 1994, assigned to the California Institute of Technology.

- 5.5 W. B. Bridges, F. T. Sheehy, and J. H. Schaffner, "60 GHz and 94 GHz Antenna-Coupled LiNbO₃ Electrooptic Modulators," IEEE Photonic Technology Letters, vol. 5, p.p. 307-310, March 1993.
- 5.6 T. Moreno, Microwave Transmission Design Data, Artech House, 1989, pp 53-55.
- 5.7 W. B. Bridges, U. V. Cummings, F. T. Sheehy, and J. H. Schaffner, "Wave-Coupled LiNbO₃ Directional Coupler Modulators," Paper 4.3, Photonic Systems for Antenna Applications (PSAA-IV), Monterey, CA, 18-21 January 1994.
- 5.8 L. J. Burrows and W. B. Bridges, "Slot-Vee Antenna Coupled Electro-Optic Modulator," Proceedings of the SPIE Conference on Photonics and Radio Frequency II, San Diego, California, July 1998, SPIE Vol. 3463, p.p. 56-65

6.0 Drift in the Characteristics of Integrated Optical Waveguide Modulators.

6.1 Background.

A serious practical problem in integrated optical waveguide modulators fabricated in lithium niobate is the drift in their characteristics, particularly the operating bias point, with time. As we have seen in Section 4, the harmonic generation and intermodulation distortion is particularly sensitive to the static operating point on the modulator's transfer function. This point is usually set by exact fabrication of the waveguide lengths (and coupling in the case of the directional coupler modulator), trimming the phase shift in the fabricated guides by laser ablation, or by applying a dc bias voltage. Subsequent drift in this operating point may be due to several different physical effects:

- (1) Temperature gradients that cause the phase shift in one part of the modulator to change differently than the phase shift in another part.
- (2) Overall temperature changes that cause changes in the coupling in the case of the directional coupler modulator.
- (3) Drift occurring with the application of a constant dc bias voltage.
- (4) "Aging", that is, drift that occurs with time, even if no voltage is applied.

The drift with environmental temperature and temperature gradient are both well known, and the measures required to minimize their impact are also well known. Our studies have been directed to try to understand the drift that occurs otherwise, with or without applied dc voltages. These effects are not well understood, although there are many "prescriptions" both published and unpublished that are used to minimize them (for example, "annealing" the modulator at different steps in its fabrication). There have been many claims of "low-drift modulators" in the commercial world, but these have failed to stand up to careful measurement by disinterested parties (for example, the modulator group at MIT/Lincoln Laboratories). There are even many patents directed at processes to minimize or eliminate drift, for example the extensive work of the Nagata group [Refs. 6.1,6.2]. However, the basic physics of the phenomena are poorly characterized in the literature, so we undertook to start our investigation by trying to make sense of the basics.

The work described below introduces our “model” of what is going on in integrated lithium niobate modulators from a basic physics standpoint. The “villain” in all cases turns out to be the proton, introduced intentionally or unintentionally into the crystal structure. We describe below the experiments we undertook to demonstrate this thesis, and the new techniques we developed to minimize the drift. Several applications for U. S. Patents have been filed by Caltech.

This work was initiated in the “funding gap” era, and was continued briefly under the present contract. However, additional funds to look at the drift problem specifically became available under a contract with MIT/Lincoln Laboratories [Ref. 6.3], so most of what is reported here was accomplished under that contract. The several U. S. Patents that have been applied for, if they are issued to Caltech, will give the U. S. Government the same rights as if they had been filed under this contract instead of the MIT/Lincoln Laboratories contract.

6.2 Review of Waveguide Fabrication Techniques

Chapter 7 of the previous contract’s final report [Ref. 6.4] described how waveguides are made in lithium niobate crystals by various techniques, including titanium in-diffusion and by annealed proton exchange. These techniques are reviewed briefly here for ready comparison with our new fabrication methods.

6.2.1 Titanium In-Diffused Optical Waveguides.

A thin strip of titanium is evaporated on the top surface of the lithium niobate crystal. The width of the strip is the desired width of the final waveguide, and the thickness of the metal deposited is related to (but not proportional to) the desired depth of the final waveguide. The crystal is then heated to about 1100 C for several hours, and the titanium atoms diffuse down into the lithium niobate, going into the crystal lattice both interstitially and by replacing niobium atoms. The resulting region is transparent, with both ordinary and extraordinary refractive indices increased. Typically, $\Delta n_o \leq 0.02$, and $\Delta n_e \leq 0.04$. During the diffusion process, wet oxygen (oxygen gas bubbled through water) flows over the lithium niobate wafer to retard out-diffusion of lithium oxide, Li_2O . As explained in Ref. 6.4, such loss of lithium oxide from the surface layer also produces a guiding surface, but one without lateral boundaries, an undesirable result.

Once the optical waveguides are formed, various other process steps are used to form the electrical structures. These typically include deposition (usually by sputtering) of a layer of

silicon dioxide, then evaporation of a metallic layer through a mask to form the electrodes. Additional steps of up-plating the electrodes, ion etching the silicon dioxide, etc. may also be taken. It is common practice to anneal the wafer at elevated temperature (typically in the 200-400 C range) for varying lengths of time at various steps, "to relieve stress," or for other claimed benefits. Such annealing is carried out in a wet oxygen atmosphere to prevent lithium oxide out-diffusion.

During the above processes, the electro-optic effect is not significantly changed, so that the resulting waveguide modulator has the same electro-optic properties as bulk lithium niobate.

6.2.2 Annealed Proton-Exchange Waveguides

A mask defining the location of the desired optical waveguides is deposited on the surface of the lithium niobate crystal, with openings located where waveguides are desired. Typically, this mask is made of a not-too-reactive metal such as nickel, or of silicon dioxide. The lithium niobate wafer is then immersed in a hot weak acid, such as benzoic or citric, at a modest temperature, typically 200 C for the order of minutes. Strong acids, such as sulfuric, may also be used, as well as much "weaker" sources of protons such as ordinary hot water. The protons from the acid enter the surface of the lithium niobate crystal very rapidly, so that immersion times must be controlled rather carefully. The protons replace the lithium atoms in the surface of the crystal to form a layer of $\text{Li}_{(1-x)}\text{H}_x\text{NbO}_3$. Typically, x will be the order of 0.5 (half the lithium atoms replaced by protons) in only a few minutes. Very large increases in refractive index can be realized for the extraordinary index, $\Delta n_e = 0.12$, but the change in the ordinary index is small and negative (non-guiding) $\Delta n_o = -0.04$. Thus proton-exchange waveguides only support TM modes for z-cut crystals or TE modes for x-cut crystals.

When protons are exchanged for lithium ions, there is more than one structural phase possible for the crystal lattice. For $x \leq 0.12$, the only structure possible is termed the " α " phase. This crystal structure has the same low optical loss and good electro-optical coefficients as pure lithium niobate. For $0.12 \leq x \leq 0.55$, an additional phase, the " β " phase can coexist. This crystal structure has more optical loss and reduced electro-optical coefficients. For $x \geq 0.55$, additional phases β' , β'' , β''' , ... can exist, and these are thought to have added optical loss and reduced electro-optical coefficients as well (although no measurements are known). At $x = 1$, the resulting HNbO_3 crystal has a cubic structure and has no electro-optical effect.

Clearly, it is desirable to have $x \leq 0.12$ to have low-loss, electro-optically active waveguides. But as noted above, even a few minutes contact with an acid will produce x of the order of 0.5. The solution to this dilemma that is usually adopted is to anneal the wafer after acid exposure to cause the protons to move further into the lithium niobate and occupy a larger volume, thus reducing x over the volume of the desired waveguide. Typical annealing temperatures are 300 to 400 C, again performed in a wet oxygen atmosphere to prevent lithium oxide loss. It is assumed in this annealing process that the β phase converts uniformly to the α phase, and that the entire waveguide becomes α phase when x is at or below 0.12. Unfortunately, the $\alpha \leftrightarrow \beta$ phase transition is of second order, which means that no energy is required for it to happen. Thus, it is possible that mixtures of α and β phase regions can exist for x near 0.12, and these regions can then grow or decrease in size with time or applied electric field (i.e., the bias voltage) as protons "move around" at room temperature.

6.3 Caltech Proton Exchange Waveguide Studies

At the very end of the previous contract [Ref. 6.4], we fabricated a few annealed proton exchange (APE) straight waveguides in lithium niobate, using molten benzoic acid and scrap pieces of lithium niobate. The optical transmission measurements indicated losses about the same as titanium in-diffused waveguides, which encouraged us to continue the experiments with a better mask for straight waveguides. The mask was not received until the end of the previous contract, but the work was continued (slowly) in the "funding gap" period.

In reviewing over 50 references on APE waveguides, we were struck by the varying processes used, the results obtained, and the theoretical diffusion models that were used to explain the results. As a result of this literature search and review, we undertook to develop a model of proton diffusion in lithium niobate that would predict diffusion depths accurately across the important diffusion process parameters of time, temperature and diffusing acid pH. We wished to explain not only our own experimental results, but also those already in the literature.

6.3.1 Proton Diffusion Model

In the "Standard" proton diffusion model, an Arrhenius plot (an assumed exponential variation of depth with annealing time and the initial proton density) is used to fit the parameters in the theory to the experimentally measured diffusion depth. The parameters fitted are the activation

energy E_A and the temperature-independent diffusion coefficient D_0 for protons in lithium niobate. According to this "Standard" model, both E_A and D_0 should be the same for all authors, since they describe the fundamental behavior of the diffusion process. However, the literature search clearly indicated that different authors, using different acids and different diffusion times, all obtained different values for these parameters.

In our approach, we developed a different model. Returning to the basic physics of diffusion, we note that E_A , the activation energy, is a measure of the potential barrier to making a "jump" of the proton from one site to another, and this is a fundamental property of the crystal itself (lithium niobate), which has a "characteristic length" between sites and a characteristic "jump frequency." This parameter should be the same for everybody's experiments, just as in the standard model. Then we note that the D_0 is a measure of the diffusion "driving force," which is determined by the chemical potential difference between the diffusing acid and the lithium niobate crystal. This coefficient will not be the same for experimenter's using different acids with different pH values. Thus, in our model, we fit a single value of E_A for all experiments, and allow D_0 to vary depending on the pH of the acid used. The experimental data is still fit to an Arrhenius plot, but with the new restriction that all E_A will be the same for data from all experimenters, but D_0 will depend on the acid pH.

A comparison is given in Table 6.1 for the results from three different published studies [References 6c-6e]. In the table, the depth measured by each group of experimenters

Table 6.1
Comparison of Experimental and Theoretical Diffusion Depths

Ref.	Acid	E_A eV	D_0	T °C	t hr	Meas. μm	Calc. (Ref.) μm	Calc. (ours) μm
6.5	Benzoic	0.94	1.84e9	180	4.0	0.651	1.01	0.675
"	" "	"	"	200	"	1.13	1.68	1.12
"	" "	"	"	220	"	1.82	2.69	1.80
6.6	Benzoic	0.95	3.47e9	150	1.67	0.162	0.333	0.185
"	" "	"	"	160	"	0.226	0.450	0.249
"	" "	"	"	180	"	0.399	0.790	0.436
"	" "	"	"	200	"	0.64	1.32	0.725
"	" "	"	"	220	"	1.06	2.12	1.16
6.7	Phosphoric	0.83	0.643e9	170	1.0	0.298	0.975	0.291
"	" "	"	"	190	"	0.495	1.56	0.495
"	" "	"	"	210	"	0.861	2.40	0.807
"	" "	"	"	230	"	1.24	3.56	1.27
"	" "	"	"	250	"	1.78	5.13	1.92

is compared to the depth calculation made by the same group (using the “standard” model) and by our model. Clearly, our model is a much better fit to the experimental diffusion depth than the model used by the original authors.

We have made this same comparison with many other sets of data from various references, and in every case our model gives a better match to the experimental data than that used by the experimenters themselves. Thus we conclude we now have a model that can accurately predict the diffusion depth of an optical waveguide and be scaled for use with different sources of protons.

6.3.2 Proton Mobility Measurements

We wished to measure the diffusion coefficient for deuterons in lithium niobate by a more direct process than an Arrhenius plot of waveguide depth. Fortunately, the instrument to do the measurements, a four-point conductivity probe in a controlled temperature furnace, was available in another research group here at Caltech, that of Professor Sossina M. Haile in the Material Science Department. Mr. Jonathan Burrows, a Caltech senior, made the laboratory measurements. Conductivity in such a good insulator as lithium niobate is a direct measure of the presence and mobility of charge carriers, in this case protons.

Samples of lithium niobate were prepared with a “standard” annealed proton exchange layer on one surface. Additional samples were prepared using only the wet annealing steps, that is, *without* immersion in an acid. The conductivity of these samples and a sample of “virgin” lithium niobate, as received from the manufacturer, were measured with the four-point probe set up. The results are given in Fig 6.1 as a function of $1000/\text{Temp [K]}$. (The scale on the abscissa corresponds to about 910 K at the left end and 625 K at the right end of the graph). Note that the conductivity of the APE sample has been increased by almost a factor of 2 over “virgin” lithium niobate at the higher temperatures, indicating that the protons can move as charge carriers. However, the most striking feature of this figure is that the sample prepared by annealing only, also has an increase in conductivity, and is virtually identically to the APE sample at lower temperatures. From this we would conclude that the wet annealing process introduces nearly as many protons as the full APE process

We may wonder what causes the conductivity in the “virgin” sample. Unfortunately, we do not know the answer, since we do not know what processes were used on the lithium niobate wafers by the manufacturer, such as washing or etching. In light of our work, it would not be surprising if such “virgin” pieces had a layer of protons exchanged over the surface. We had hoped to run additional studies on “virgin” pieces to see if the conductivity could be decreased by subsequent processing, indicating removal of protons, but we did not have the time.

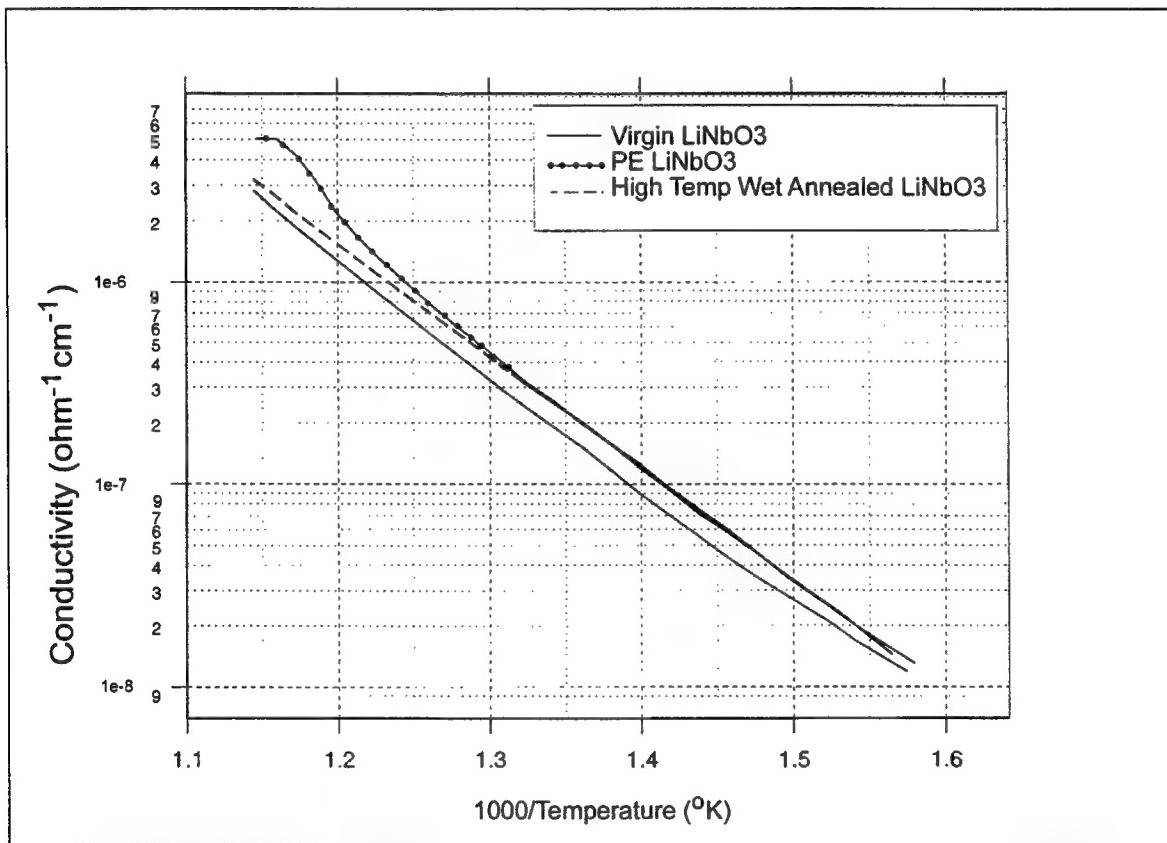


Fig. 6.1 Conductivity of lithium niobate samples as a function of reciprocal temperature.

6.3.3 Deuteron Exchange Waveguides

As a test of our model, we decided to do experiments with deuterons in place of protons, to see if the model described above would successfully predict the waveguide depths. We anticipated that the diffusion coefficient for deuterons would be significantly lower than protons. The most direct measurement would have been to repeat the four-point probe conductivity measurements

of the previous section for annealed deuterium exchange layers, and we had planned to do this in the future. However, we decided instead to see if we could actually make ADE waveguides and a modulator. As a practical matter, we would expect that any drift in the waveguide characteristics for an annealed deuteron exchanged waveguide (ADE) would be smaller than that in an equivalent APE waveguide because of the factor of two larger mass, and this was the attraction of making waveguides prior to the confirming science of the four-point probe measurements.

As far as we know from our literature search, no one has ever made an ADE waveguide. Thus, at the outset we were faced with several questions:

- * What is the prescription for ADE? We can obtain a partial answer from our diffusion model, but the subsequent annealing process would have to be determined experimentally.
- * What magnitude change of refractive index would be expected? Would it be enough (and of the correct sign) to form a waveguide?
- * Would the resulting waveguide have high optical quality (low loss)?
- * Would the resulting waveguide still exhibit an electro-optic effect?
- * What should we use a source of deuterons?

Making an ADE modulator was clearly a step into the unknown! The last question above was obviously the first that had to be answered in an experimental program. While APE waveguides can be made from many readily obtainable acids (including water), a pure source of deuterons is more challenging (and expensive). We considered using heavy water, D₂O, but we weren't certain that our diffusion model would be scaleable that far away from conventional acids. Since benzoic acid, C₆H₅COOH, is the most popular acid source for APE, that would have been the logical choice. Unfortunately, the fully-deuterated equivalent, C₆D₅COOD, is very expensive and obtainable on special order only. We decided to use a less expensive and more easily obtainable source instead, deuterated sulfuric acid, D₂SO₄. Sulfuric acid has a much lower pH than benzoic acid, making it a more copious source of protons (or deuterons) than benzoic acid. But we anticipated that we would be able to correct for the different pH with our new model.

Working with concentrated D₂SO₄ in a diffusion furnace is also considerably more dangerous than with molten benzoic acid. Because of the lower pH, the initial diffusion times are much shorter. With benzoic acid, the vessel containing the acid and lithium niobate chip can be loaded into the furnace and the temperature ramped up, held, and ramped down. With the lower pH for sulfuric acid, the times are shortened to the point that the furnace will not ramp up or down sufficiently fast to obtain the correct diffusion time. Thus, the lithium niobate chip has to be lowered into the sulfuric acid while it is in the furnace at the correct temperature. And the (pure, fuming) sulfuric acid must be contained in a closed, vented vessel to prevent destruction of the furnace. Thus we must "build a ship in a bottle" by lowering and raising the lithium niobate chip quickly into the acid, inside a closed vessel inside the hot furnace.

The initial experiments were done with a mask for a simple straight waveguide section. The diffusion time was corrected using our theory, but the subsequent annealing time was more or less a guess at this point. The result was a waveguide that was single mode but lossy and clearly too shallow. The next sample was annealed for a substantially longer time. The resulting waveguide was still single mode and exhibited excellent optical properties. The measured loss was actually slightly better than our best APE guides.

Encouraged that we could make optically good waveguides, we next used a simple Mach-Zehnder modulator mask. We used the diffusion times and anneal times from the straight waveguide experiments. We deposited simple M-Z electrodes on the modulator. The results was a V_π of about 10 volts, comparable to our APE and titanium in-diffusion modulators made with the same mask. The optical transmission was also comparable or better. Thus we conclude that we could make successful modulators via ADE rather than APE, (perhaps with even better optical qualities), and these would presumably have lower bias drift or aging characteristics (possibly lower temperature sensitivity) than APE modulators because of the heavier, larger deuterons. Unfortunately, the contract ended before we were able to make actual drift comparison measurements, as described below.

6.3.4 APE Modulator Drift Measurements

In order to have baseline values for drift in APE modulators, we fabricated a simple Mach-Zehnder modulators and set it up to measure its bias drift rate. The V_π for this modulator was the order of 10 V.

The drift measurements were made in the following fashion: The dc bias and a small sinusoidal signal at 500 Hz were summed and fed to the modulator electrodes. The resulting optical output was detected and fed to a lock-in amplifier tuned to 1000 Hz. (The 500 Hz sub-harmonic is actually obtained from the lock-in 1000 Hz reference oscillator). The lock-in output is low-pass filtered and represents the amplitude of the second harmonic distortion, and thus the departure from the correct bias point. This signal is amplified by an operational amplifier, and then added to the bias signal. The manual bias is used to set the modulator initially to the correct bias point, which should produce zero second harmonic of the 500 Hz signal. As the modulator drifts off this point, a 1000 Hz signal develops. The level and sense of this signal thus provide an error, which is fed back to the modulator with high gain to maintain a near-zero output of the 1000 Hz signal. The amplitude of the error signal applied to the modulator is thus a direct measure of the drift of the bias point in equivalent volts.

Figure 6.2 shows the resulting bias drift with time, plotted as $\Delta V/V$, where ΔV is the bias error in Volts and V is the initial bias applied to the modulator. In Fig. 6.2, the initial bias was 7.6V. We note that $\Delta V/V$ is about 0.2 in about 24 hours, so the bias drift is about $0.2 \times 7.6 = 1.5$ V in 24 hours. The time it takes to make $\Delta V/V = 1$ is used in the telecommunications industry to judge the acceptability of a modulator due to bias drift. For a modulator for critical applications, the industry target time is about 25 years! The extrapolated value for our modulator is 9.4 days; clearly, our APE modulator would be unacceptable.

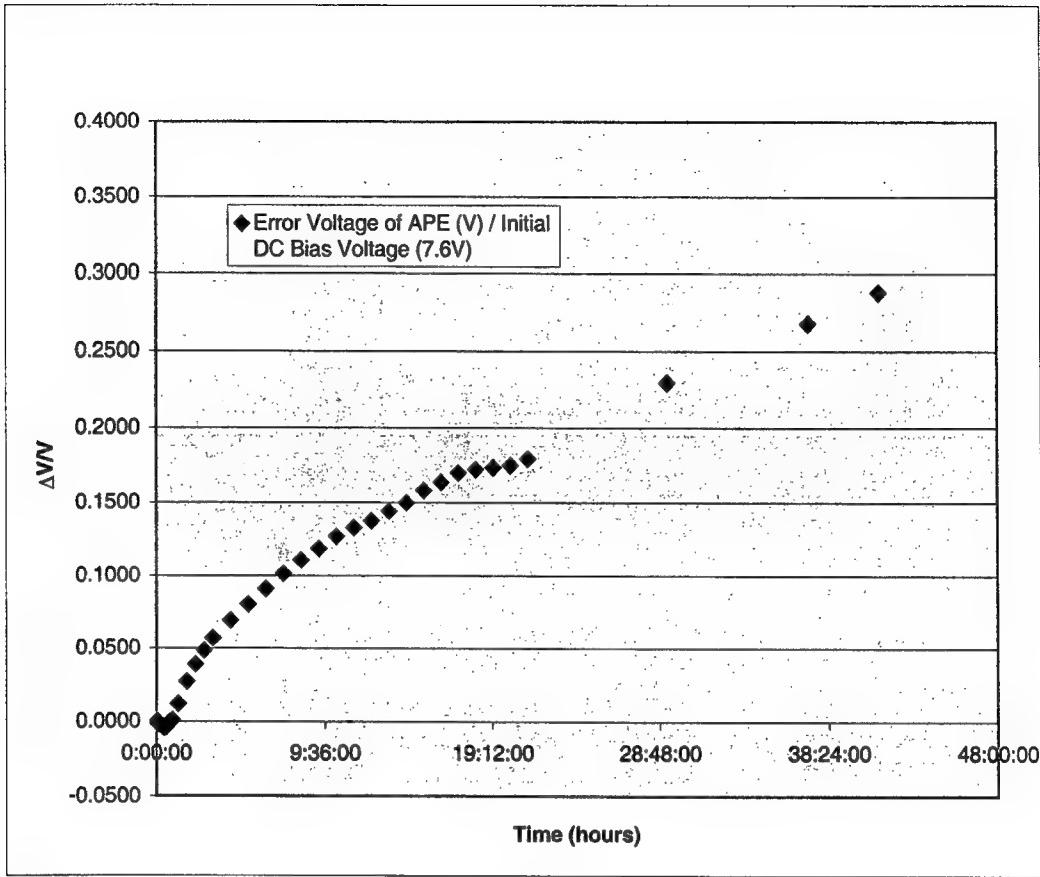


Figure 6.2 Bias voltage drift referred to the initial bias voltage of 7.6 V. for a APE MZ modulator. The time to $\Delta V/V=1$ is extrapolated to 9.4 days

Figure 6.3 shows the results for the same modulator, but with the initial bias set to a substantially higher value, 52 Volts. (This is a couple of cycles of V_π higher than the "normal" operating point). Note that the drift rate measured as $\Delta V/V$ is higher; we reach $\Delta V/V = 0.2$ in about 9.5 hours, or about 2.5 times faster than the test with lower bias voltage. But remember that the initial voltage is also higher, so that the change measured in volts is now $0.2 \times 52 = 10.4$ volts. Thus we could say that the drift in actual volts in this case is $2.5 \times (10.4/1.5) = 17.3$ times faster than with the lower initial bias voltage. It is probably better to make this comparison using the long-term drift trends. Thus we compare the extrapolated values: $(9.4 \text{ days}/1.82 \text{ days}) \times (52 \text{ V}/7.6 \text{ V}) = 35.3$ times higher drift rate in volts at the higher initial bias.

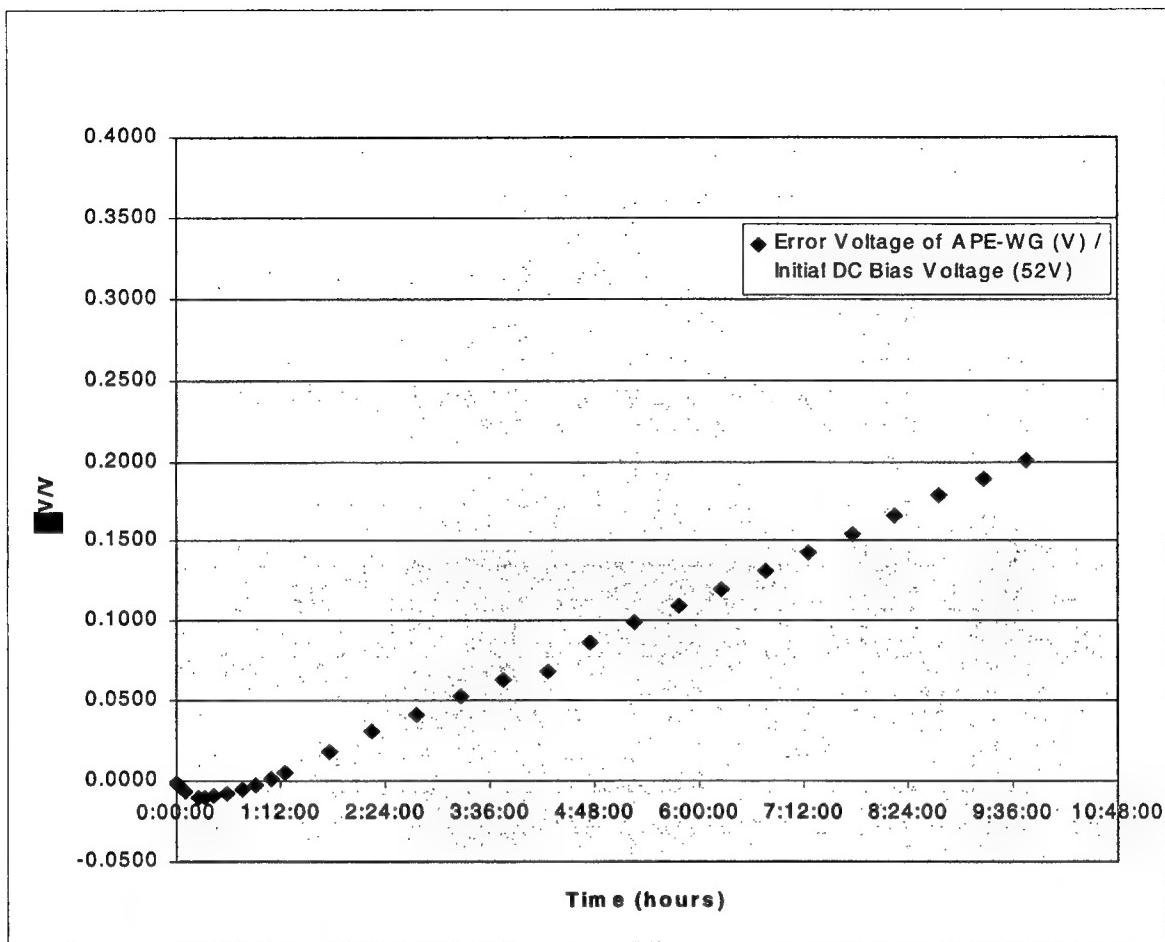


Figure 6.3 Fractional drift for the same APE modulator as Fig. 6.2, but with an initial bias of 52 V. The extrapolated time to $\Delta V/V=1$ is 1.8 days.

6.3.5 Conclusions From Proton Exchange Studies

We may conclude several things from our proton exchange studies:

- * We have an excellent, scalable model for diffusion of protons (and deuterons) into lithium niobate.
- * We can make modulators with ADE that are as good or better optically and electrically than the equivalent APE modulators, and they should exhibit better bias drift and aging properties than APE modulators.
- * We are also convinced, based on the above studies, that bias drift and aging in APE and ADE modulators can never be eliminated, since they result from the mobility of protons (or deuterons) in the crystal.

While this last conclusion may be discouraging for APE and ADE modulators, we also concluded the following:

- * If we can exclude protons from modulators made by a different process, perhaps we can eliminate bias drift and aging from those modulators.

And that turned our research on drift and aging into another direction, as described in the next section.

6.4 Proton-Free Titanium In-Diffused Modulators

6.4.1 Drift and Aging in Standard Titanium In-Diffused Modulators

Titanium in-diffused modulators are known to exhibit bias drift and aging, just like APE modulators, although generally not so pronounced nor so consistent in properties. Various causes of these undesirable properties have been proposed and “solved” by patented cures, for example, Refs 6.1 and 6.2. While we cannot exclude these causes from our studies, we certainly can add another: protons in the lithium niobate itself. We are convinced by our proton diffusion data that protons can easily enter the lithium niobate substrate at many steps in Ti modulator fabrication and result in drift and aging of the modulator. Any washing or cleaning step, for example, is a potential source of proton exchange. But the most obvious source of protons is the “wet oxygen atmosphere” that is used in every furnace step: The in-diffusion step and subsequent annealing steps. As we stated in section 6.2, the wet oxygen atmosphere is thought to be essential to prevent the loss of lithium oxide at high temperatures, so it is universally used. It is “well known” that using dry oxygen by simply omitting the water “bubbler” does not work. The resulting chip surface is usually milky, and the optical losses are unacceptable. The milky appearance is attributed to the conversion of the surface layer to LiNb_3O_8 , a material that is not electro-optic.

6.4.2 Titanium In-Diffused Modulator Drift Experiment

In order to provide baseline drift measurements for our studies, we fabricated titanium in-diffused Mach-Zehnder modulators according to the conventional wisdom, as described in Sec. 6.2.1. These modulators exhibited good optical properties and a V_π of the order of 10 V, similar to the APE and ADE made with the same mask. One of these modulators was set up in the drift

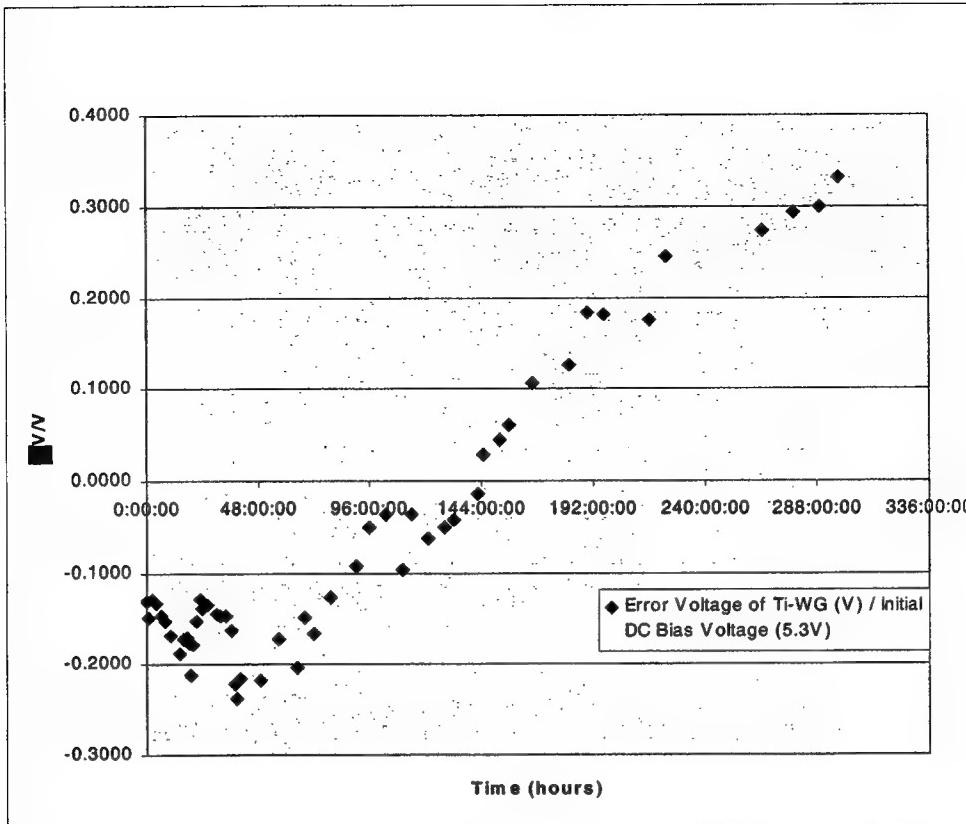


Figure 6.4 Drift in bias voltage relative to an initial bias of 5.3 V for a titanium in-diffused modulator made by conventional techniques. The extrapolated time to reach $\Delta V/V=1$ is 19.5 days

test measurement setup and run. Again, the bias point was set manually to the operating point for zero second harmonic distortion, 5.3 V in this case. Figure 6.4 shows the subsequent drift. The projected time to reach $\Delta V/V=1$ is 19.5 days, better than the APE modulator, but still clearly unacceptable for commercial applications. At higher bias voltages the drift is much worse, just as it was for APE modulators. Figure 6.5 shows the drift of the same modulator when the initial bias was reset to 51 V, another point giving zero second harmonic distortion. The drift rate is substantially faster. The extrapolated time to reach $\Delta V/V=1$ is 3.6 days. Using these extrapolated values, the drift rate in volts is $(19.5/3.6) \times (51/5.3) = 52$ times as fast for the higher initial bias voltage. Clearly, this titanium in-diffused modulator would be unsatisfactory for most applications.

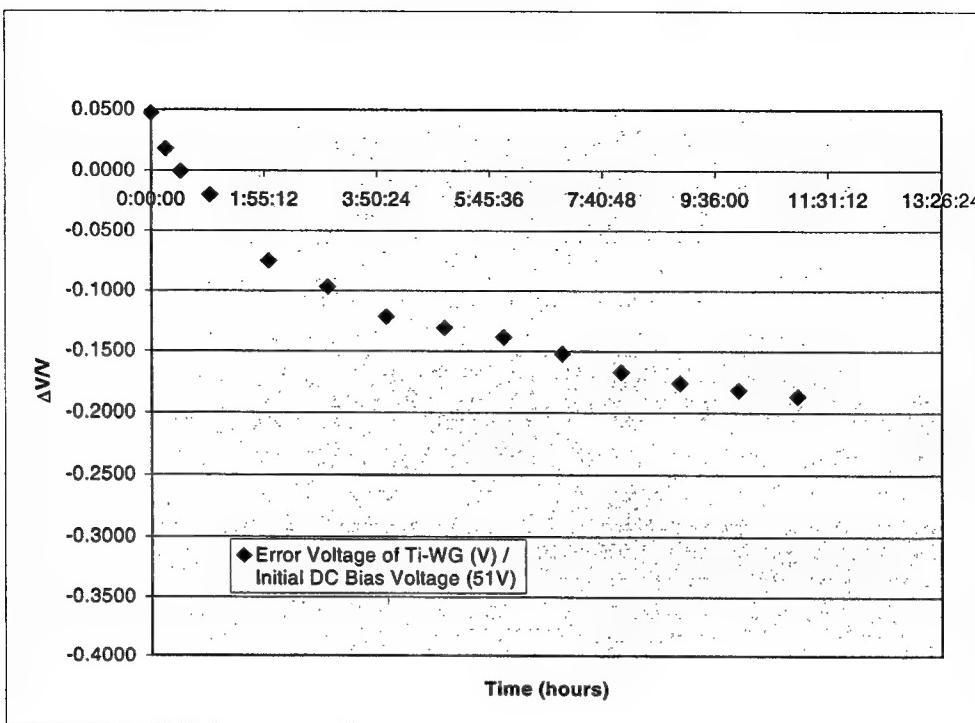


Figure 6.5 Bias drift in the same modulator as Fig. 6.4, but with initial bias of 51V. The extrapolated time to reach $\Delta V/V=1$ is 3.58 days.

6.4.3 Discovery of a Proton-Free Process

Armed with the thesis that drift in titanium in-diffused modulators was caused by protons “unintentionally” introduced during the fabrication process, we set about to discover a process that would yield good modulators, but without the “wet oxygen” used during all annealing steps and without acids and hydrogen-containing solvents in the cleaning steps.

We set out to convince ourselves that the “common wisdom” was correct. First, we processed lithium niobate samples in dry oxygen atmospheres. Indeed, we obtained milky looking samples. Next, we pursued the additional “common wisdom” that providing “sacrificial” samples of lithium niobate near the modulator chip being processed in the furnace would slow the out-diffusion of lithium oxide. The theory is that the gas, flowing over these sacrificial samples, picks up lithium oxide, and provides a “lithium oxide rich” atmosphere over the desired modulator chip, thus preventing the oxide out-diffusion. This worked with flowing wet oxygen, but the end result was no better nor worse than the process without the sacrificial samples. We

also tried dry oxygen with sacrificial samples, but this resulted in milky modulator chips as before. We concluded that the "sacrificial sample" theory was simply not working, at least as we had implemented it.

An essential discovery came when we reasoned that perhaps the "sacrificial sample" idea was correct, but that it needed a significant quantitative change to make it actually work. We reasoned that two things could be improved: (1) A greatly increased surface area of the sacrificial lithium niobate compared to the lithium niobate modulator surface, and (2) the use of pressurized oxygen to enhance the pickup rate of lithium niobate from the sacrificial material. We fabricated a closed stainless steel pressure vessel for the furnace, and loaded it with powdered lithium niobate as well as a lithium niobate test chip. The assembly was raised to 1100 C to simulate the in-diffusion process. A static (non-flowing) dry oxygen pressure of a few p.s.i. was maintained during the process. The results of this test were mixed: To our delight, the test chip remained clear. But it was covered with baked-in particles of lithium niobate powder. Evidently, the pump down pressure transient and oxygen fill process had stirred up the lithium niobate powder and scattered it all over the insides of the chamber.

Encouraged by the results, we designed and fabricated a ceramic inner chamber with porous end plugs which would allow the entry and exit of oxygen, but were fine enough to block any lithium niobate powder that would be stirred up during the pump-down and oxygen fill process. The next run produced the desired result: a clear-looking lithium niobate chip that had been cycled to 1100 C in dry oxygen.

There were now several avenues to pursue: We made several runs as above at various pressures of oxygen from 2 to 200 p.s.i., to determine if there was any change in performance. It was clear that a few psi pressure above ambient was all that was necessary, and that high pressures were undesirable from a safety standpoint (200 p.s.i. of oxygen at 1100 C degrades the stainless pressure vessel rather rapidly, besides having "bomb-like" potential). Having determined an "optimum" pressure, we next determined that oxygen was in fact necessary. Runs were made with argon, and nitrogen at above-ambient pressures. The results confirm that it is only oxygen that leaves the lithium niobate samples in the desired form. Argon turned the samples a nice amber color. Nitrogen turned them jet black!

6.4.4 Driftless Titanium In-Diffused Modulators

The next step was to make waveguides and modulators using this process. The first step was to evaporate a uniform titanium layer on a lithium niobate wafer. Next, a sputtered silicon dioxide waveguide mask was then laid down, and the titanium layer was then argon ion etched through the mask to leave the desired titanium stripes on the substrate. The titanium was in-diffused at the standard temperature, but now inside the stainless steel pressure vessel and the ceramic cylinder with porous end plugs at a few psi of oxygen. After the diffusion, gold electrodes were applied to form the modulator electrodes.

The optical tests run on the straight waveguide sections exhibited optical transmission as good or better than conventional titanium in-diffused waveguides. Unfortunately, too much titanium had been used in the initial layer, so the resulting waveguides were multimode, and the resulting modulator had a poor extinction ratio (only 3 dB or so). Despite the reduced slope caused by the poor extinction ration, the sensitivity was essentially the same as the conventional titanium in-diffused modulator (which was also multimode, having come from the same batch of “too thick” titanium evaporation) and comparable to the, APE and ADE modulators, all made with the same mask.

However, the most amazing feature of this modulator became apparent when it was run in the bias drift setup. Figure 6.6 shows the measured drift in the applied bias necessary to maintain the modulator at a zero second harmonic operating point with a relatively high bias voltage of 56 V. Shown for comparison are the previous APE and “conventional” titanium in-diffusion modulator drift results from Sec 6.3.4 and 6.4.2. There appears to be no drift trend! (The extrapolated value of the time to reach $\Delta V/V=1$ is 457 years, but such an extrapolation is unwarranted based on only 100 hours of data). There appears to be a periodic variation in the bias value with a period of 24 hours. An expanded scale version of the bias drift is shown in Fig. 6.7. We conclude that this periodic variation is due to the change in environmental temperature due to the cycling of the laboratory air conditioning system. In our measurement setup, the modulator chip is out in the open air, with a butt-coupled fiber input (no physical contact) and a lens coupled output.

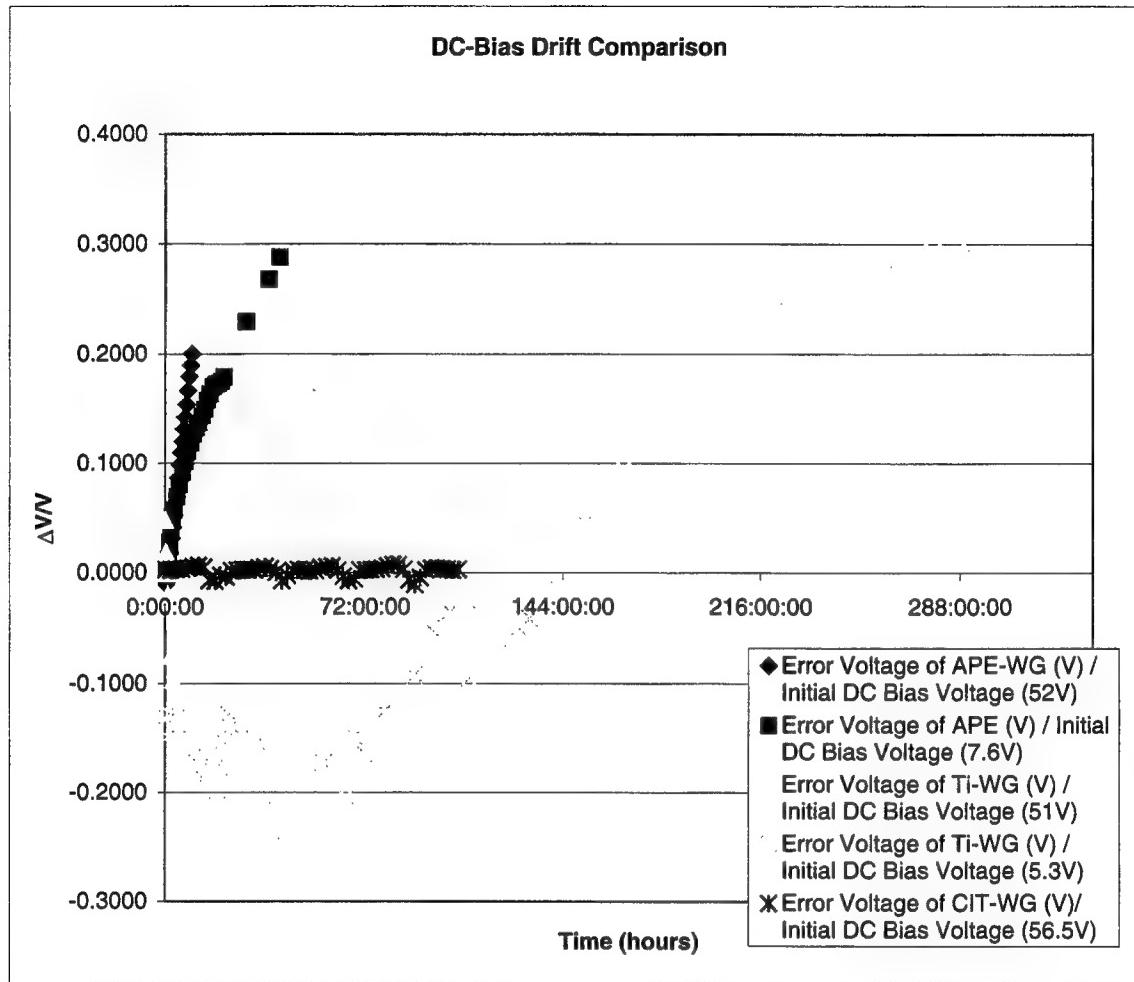


Figure 6.6 Drift rate for the modulator made by our new dry oxygen process compared to the drifts for APE and conventional Ti in-diffused modulators. The extrapolated value of the time to reach $\Delta V/V=1$ is 457 years

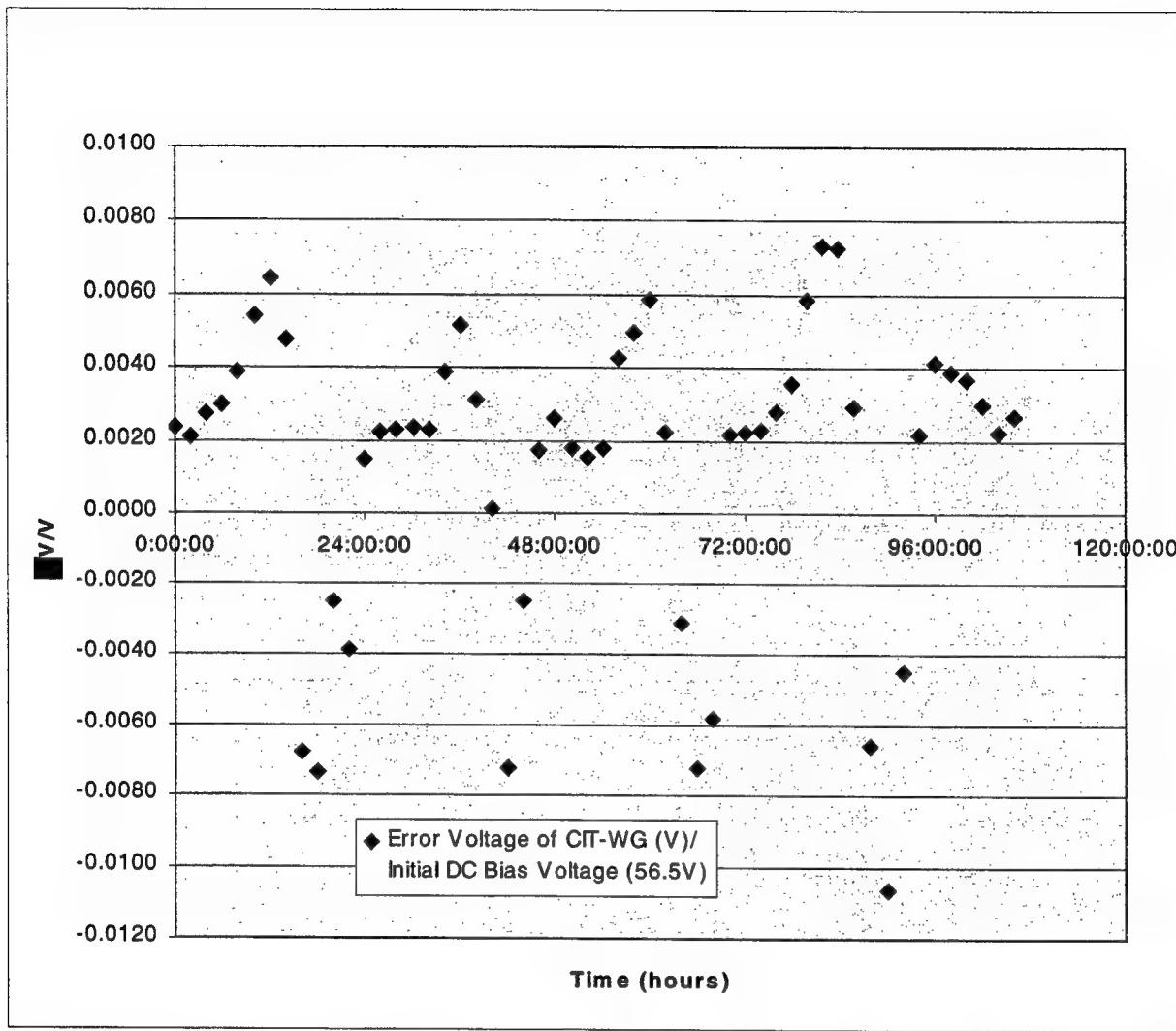


Figure 6.7 Bias drift in the Caltech process modulator of Figure 6.6 on a greatly magnified scale. The periodic variation is due to room temperature changes.

Unfortunately, the actual modulator fabrication and the drift measurements were made substantially after the expiration of both the present contract and the contract with Lincoln Laboratories. The student who developed the processes to make this modulator, Lee J. Burrows, left Caltech to take a position with the New Focus Corporation shortly after making these measurements. Since the Principal Investigator for the project, Prof. Bridges has committed to retire in June 2002, there were no new graduate students in the pipeline to continue this work. Thus, there was no opportunity to make a second modulator to confirm the apparently near-zero drift features. For that reason, these results must be taken as tentative until they have been confirmed by further experimental measurements. Patent disclosures were written for this process, however, and were filed by Caltech. They are now pending.

6.4.5 Future Plans for Low Drift Modulators

No further funding has been sought specifically for low-drift modulator studies at Caltech, primarily because Prof. Bridges anticipates retiring in June 2002. However, Mr. Lee J. Burrows, who was responsible for these studies at Caltech is now Director of Modulator Development at New Focus, Inc., and will be likely continue to pursue the low drift techniques.

6.5 Conclusions

The studies of the basic material properties of lithium niobate stimulated by the nagging problems of drift and aging in modulators have been fruitful, and resulted in a better understanding of protons in lithium niobate. Practical results include the first demonstration of annealed deuterium exchange (ADE) waveguides and modulators, with promising properties compared to conventional APE waveguides, and the discovery of processes to make titanium in-diffused waveguides that appear to be proton-free and drift-free.

6.6 References

- 6.1 H. Nagata, J. Minowa, M. Yamada, J. Ichikawa, J. Nayyer, S. Shimotsu, J. Ogiwara, M. Kobayashi, H. Honda, K. Kiuchi, T. Shinriki, and T. Sakamoto, "Optical waveguide modulator having a reduced d.c. drift," U. S. Patent No. 5,526,448, June, 1996.
- 6.2 H. Nagata, M. Shiroishi, T. Kitanobou, and K. Ogura, "Dc drift reduction in LiNbO₃ optical modulators by decreasing the water content of vacuum evaporation deposited SiO₂ buffer layers," Optical Engineering, Vol. 37, p.p. 2855-2858, October 1998.
- 6.3 L. J. Burrows and W. B. Bridges, "Low Drift Optical Waveguide Fabrication in Lithium Niobate," MIT Purchase Order BX-5913 under Air Force Contract F19628-95-C-0002, CHANGE ORDER 5.
- 6.4 W. B. Bridges, L. J. Burrows, U. V. Cummings, R. E. Johnson, and F. T. Sheehy, "60 and 94 GHz Wave-Coupled Electro-Optic Modulators," Final Technical Report RL-TR-96-188 on contract F30602-92-C-0005 from Rome Laboratory, USAF, September 1996.
- 6.5 D. F. Clark, et al., J. Appl Physics, Vol. 54, p.p. 6218-6220, November 1983.
- 6.6 C. Canali, et al., J. Appl. Physics, Vol. 59, p.p. 2643-2649, April 1986.
- 6.7 E. Y. B. Pun, , K. K. Loi, and P. S. Chung, "Proton-Exchanged Optical Waveguides in Z-cut LiNbO₃ Using Phosphoric Acid," J. Lightwave Technology, Vol. 11, p.p. 277-284, February 1993.

7.0 PERSONNEL

The Principal Investigator for this program is Professor William B. Bridges, who was directly involved in all phases of the work.

Mr. Lee J. Burrows received his B. S. in Applied Physics from Caltech in June 1992, and joined our research group that same month. He received his M. S. in Applied physics in June 1993, and anticipates receiving the Ph. D. degree in Applied Physics in June 2001. He has been employed as a Graduate Research Assistant on the current contract over its duration. Mr. Burrows took up the task of demonstrating of Dr. Finbar T. Sheehy's idea for the Slot-V antenna-coupled modulator (Section 5 of this report), and also has undertaken the study of physical processes that lead to drift in waveguides (Section 6 of this report). He is current Director of Modulator Development at New Focus Corporation.

Mr. Uri V. Cummings worked during the summer of 1993 under Caltech's Summer Undergraduate Research Fellowship (SURF) program, and continued as a Senior Thesis student in the academic year 1993-94 on the work begun by Dr. Sheehy on the mm-wave directional coupler. He received his B. S. in Electrical Engineering from Caltech in June 1994, and joined our research group as a Graduate Research Assistant. He received the M. S. in Electrical Engineering in June 1995 and anticipates receiving a Ph. D. in Electrical Engineering in June 2001. He has worked on this project over its duration. He continued the work on the mm-wave directional coupler (Section 3 of this report) and initiated work on the effect of velocity mismatch on linearized modulators (Section 4 of this report). He is currently one of the founders and President of Asynchronous Digital Design Corporation.

Mr. Reynold E. Johnson has contributed to the project as Staff Engineer on a part time basis since its inception. He has been with our research group at Caltech since 1988, when he retired from the Hughes Aircraft Company, Missile Group. He retired from Caltech in January 2000.

Dr. Axel Scherer, Professor of Electrical Engineering at Caltech, has been of invaluable assistance in helping us to setup our modulator and waveguide fabrication capability, and in our attempts at understanding the physical structure of lithium niobate. He has allowed us generous use of his time and equipment.

Dr. James H. Schaffner at HRL Laboratories, LLC, continued his informal collaboration with this project in its early stages, fabricating the first of the modulator test samples for us.

8.0 INTERACTIONS, PAPERS, AND PATENTS

8.1 MEETINGS AND CONFERENCES

During the contract period, Mr. Norman P. Bernstein attended briefings at Caltech on 13 February 1996, 18 June 1996, 9 September 1996, 10 January 1997, 27 June 1997, 1 August 1997, 13 November 1997, and 21 October 1998.

The following conferences were attended during the contract period by project personnel:

(Cummings) LEOS 9th Annual Meeting, Boston MA, November 1996. "Effects of Velocity Mismatch and Transit Time on Linearized Electro-Optic Modulators," paper WW3.

(Bridges) Microwave Photonics '96, Kyoto Japan, December 1996. "Antenna-Coupled Millimeter-Wave Electro-Optic Modulators for 20 to 100 GHz," paper Tu1-6; "Linearizing Technology for Microwave Photonics (Invited)," paper Tu4-2.

(Bridges, Burrows, Cummings) Seventh Annual DARPA Symposium on Photonic Systems for Antenna Applications (PSAA-7), January 1997. "Effects of Velocity Mismatch and Transit Time on Linearized Electro-Optic Modulators," paper 5-4 (presented by Cummings).

(Cummings) Eighth Annual DARPA Symposium on Photonic Systems for Antenna Applications (PSAA-8), February 1998

(Bridges, Burrows) SPIE Conference on Photonics and Radio Frequency II, San Diego CA, July 1998. "Slot-Vee Antenna-Coupled Electro-Optic Modulator," paper number 3463-07, presented by Burrows.

8.2 PUBLICATIONS

U. V. Cummings and W. B. Bridges, "Bandwidth of Linearized Electro-Optic Modulators," J. Lightwave Technology, Vol. 16, pp. 1482-1490, August 1998.

8.3 PATENTS

U. V. Cummings, "Traveling-Wave Reflective Electro-Optic Modulator," U. S. Patent 5,886,807, March 23, 1999.

GLOSSARY

Align-Rite	A commercial photolithographic mask maker in Burbank CA
Annealed Deuteron Exchange waveguides	Deuteron exchange waveguides in which the concentration of deuterium nuclei has been reduced by diffusion in depth by subsequent heating.
Annealed Proton Exchange waveguides	Proton exchange waveguides in which the concentration of hydrogen (protons) has been reduced by diffusion in depth by subsequent heating.
Antenna-coupled modulator	An electro-optical modulator wherein the electrical signal is coupled in through antennas disposed on the dielectric-air interface.
ATX	A solid state laser manufacturer..
Band designations:	
X-Band:	The frequency range is 8.2 to 12.4 GHz
K-Band:	The frequency range is 18 to 26.5 GHz
V-Band:	The frequency range is 50 to 75 GHz
W-Band:	The frequency range is 75 to 110 GHz
Bias	The dc voltage applied to an electro-optic modulator to set the correct operating point.
Bias drift	The undesirable change in operating point of a modulator with time.
Bolometer	A device to measure rf power by measuring the change of resistance of a wire or film when it is heated by that power.
Butt coupling	Coupling of an optical wave from one waveguide to another by locating the two ends very close together and aligning them to be coaxial.
BWO Backward wave oscillator..	A tunable microwave vacuum tube oscillator
Caltech	California Institute of Technology
"Canonical" parameters	The table of modulator parameters used by Bridges and Schaffner in Reference 4.2 to compare different modulators.
DCM	Directional Coupler Modulator

DCM.43	A DCM operated at its normal bias of $0.43V_s$. to minimize second harmonic distortion.
DCM.79	A DCM operated at a bias of $0.79V_s$ to null the third-order intermodulation distortion. Also denoted SDCM.
DCM2DC	A DCM with two additional bias sections that can be adjusted to minimize both second harmonic and intermodulation distortion. Also known as the "UCSD modulator" or DCM2P.
Deuteron-exchange Waveguides	Optical waveguides (here, in lithium niobate) made by exchanging deuterium for lithium in the crystal lattice.
Dicing saw	A precision saw used to cut crystalline materials very accurately and smoothly.
Dielectric image guide	A form of slab dielectric guide in which one side of the slab rests on a conductor. The modes that propagate are thus only the odd modes of the full dielectric slab waveguide, required by the "image" in the conductor.
DPMZM	Dual parallel Mach Zehnder Modulators. A linearization scheme using two MZM's in parallel optically.
DSMZM	Dual series Mach Zehnder Modulators. A linearization scheme using two MZM's in series optically.
Dynamic range	The ratio of the signal level out of a device to the intermodulation distortion out of the device when the intermodulation distortion exactly equals the noise level.
f	The modulation frequency in an electro-optic modulator.
Ferrite isolator	A waveguide device using a ferromagnetic ceramic and a dc magnetic field to provide one-way propagation through the device, using the Faraday effect.
FP/APC connector	A bayonet-type industry standard connector for single mode polarization-preserving optical fibers. APC stands for "angle polished connector," indicating the end is cut at a slight angle to reduce retro-reflections.
Free spectral range	The range, in frequency, between successive transmission peaks of a multi peak filter, especially the Fabry-Perot interferometer filter.
Friis formula	Famous formula for the propagation loss between two antennas in free space stated by Harald T. Friis of Bell Laboratories in the 1930's.

FXR	A microwave equipment manufacturer of the 1960's, now out of business.
GRIN lens	GRaded INdex lens element made by forming a radial variation in the refractive index, rather than a radial variation in thickness.
HF	A lossy, high dielectric constant material made by Emerson and Cumings.
HRL	Hughes Research Laboratories prior to 1998; HRL Laboratories LLC thereafter.
Intermodulation distortion	Undesired signals generated by the nonlinearity of a device. Usually the signal at $2f_1 - f_2$ or $2f_2 - f_1$ when signals f_1 and f_2 are applied to the device.
Klystron	A microwave vacuum tube oscillator.
L, L_1 , L_2	Lengths of the electrodes (or portions thereof) in an electro-optical modulator
Link	See photonic link.
LiNbO_3	Lithium Niobate; a transparent crystalline material which displays a strong electro-optic effect.
Linearized modulators	Electro-optic modulators in which some means of linearizing the transfer function about the operating point has been incorporated.
Lock-in amplifier	An electronic device using synchronous detection (see below).
m	Modulation fraction in amplitude modulation
mm-waves	Electromagnetic waves with wavelengths shorter than 10 mm.
"Mouse-bites"	Chips out of the edges of the metallization on a mask used to make optical waveguides in modulators.
Multimode waveguide	Optical waveguide where more than one mode can propagate if excited
MZM	Mach-Zehnder modulator; a modulator based on the Mach-Zehnder interferometer configuration
n, n_o	Optical refractive index; the ratio of the speed of light in a vacuum to the speed of light in a material.

n_{eff} , n_m	The effective value of n for a transmission system of complicated composition, e.g., the ratio of the velocity of light in vacuum to the signal velocity along a transmission line deposited on a dielectric interface.
Optical circulator	An optical device, usually three-port, in which transmission is allowed from port 1 to port 2, port 2 to port 3, and port 3 to port 1, but not in the opposite directions. Based on the Faraday effect in glass.
P	In the analysis section, a variable referring to "other" unspecified parameters
P_m	Rf modulation power into an electro-optical modulator.
Photonic link	An optical communication link consisting of a laser, a modulator, an optical fiber, and a detector.
Photoresists, positive	Photosensitive emulsion applied to a surface that hardens and remains where it is not exposed to light.
Photoresist, negative	Photosensitive emulsion applied to a surface that hardens and remains where it is exposed to light.
Protons	Hydrogen atom nucleii. Hydrogen atoms introduced into the crystal lattice are ionized and become imbedded protons.
Proton-exchange waveguides	Optical waveguides (here, in lithium niobate) made by exchanging hydrogen for lithium in the crystal lattice.
PSAA-	Photonic Systems for Antenna Applications conference; numbered sequentially
R_m	The characteristic impedance of the modulation rf electrodes; the matched termination for a modulator.
SDCM	See DCM.79.
SDSMZM	Sub-octave dual series MZM. A particular bias condition of the DSMZM that permits large second harmonic content.
SFP	Scanning Fabry Perot interferometer. An optical spectrum analyzer based on transmission through a Fabry Perot interferometer filter.
$\text{Sinc}(x)$	Abbreviation for the $\sin(x)/x$ function
Single-mode waveguide	Optical waveguide in which only the lowest order optical mode can propagate, even if higher order modes are excited initially.

Skin effect	The penetration of rf electric and magnetic fields into an imperfect conductor. These fields decay exponentially, with an e-folding depth termed the "skin depth."
Slab waveguide	A dielectric waveguide made of a simple parallel-sided slab of dielectric material.
Slot Vee Antenna	The dual of the Vee antenna, where conductor is replaced by aperture and vice versa in a metal ground plane.
SPIE	Society of Photographic Instrumentation Engineers.
Stycast®	Trademark for artificial dielectric material by Emerson and Cummins Co.
SURF	Summer Undergraduate Research Fellow (or Fellowship) at Caltech.
Synchronous detection	Detection of a periodic signal in noise by multiplication of the signal with a "clean" sample of identical period, and then time-averaging the result. The effective noise bandwidth is essentially the bandwidth of the averaging RC filter and can be made very small.
Titanium in-diffused waveguides	Optical waveguides formed by diffusing titanium atoms into a crystal lattice at high temperature.
Transfer function	The ratio of output to input optical levels of an electro-optical modulator as a function of the applied voltage.
Transit time	Here, the transit time of the optical signal or modulating signal through the modulator electrode region.
UVC-el-1	Directional coupler electrode mask designed by U. V. Cummings.
Vee Antenna	Dipole antenna with dipole arms forming an angle less than 180 degrees. Usually the dipole arms are several wavelengths long.
Velocity mismatch	The difference between the velocities of the optical signal and the electrical modulation signal through the modulator, usually expressed in terms of the difference in refractive indecies.
V_m	The peak rf modulation voltage applied to a modulator.
V_s	The cross-over voltage for a directional coupler modulator. The voltage change required to turn the modulator from maximum transmission to minimum transmission.

V_π	The Half-Wave voltage for a Mach-Zehnder modulator. The voltage change required to turn the modulator from maximum transmission to minimum transmission.		
WR-	Electronics Industry Association designation for standard waveguide:		
	WR-90 0.400" x 0.900"	8.2 -12.4 GHz	X-band
	WR-42 0.170" x 0.420"	18 - 26.5 GHz	K-band
	WR-15 0.075" x 0.150"	50 -70 GHz	V-band
	WR-10 0.050" x 0.100"	75 -110 GHz	W-band
x	Atomic fraction of an element in a crystal chemical composition.		
x	A crystallographic axis in lithium niobate.		
Y-fed DCM	A directional coupler modulator in which each arm is excited equally by an optical wave, divided by a preceding optical Y-junction.		
YFDCM	A Y-fed DCM		
z	The coordinate in the direction of propagation in an optical waveguide.		
z	A crystallographic axis in lithium niobate.		
α	Designation for the phase in lithium-hydrogen niobate crystalline system that occurs for the smallest value of the hydrogen fraction x.		
$\beta, \beta', \beta'', \beta''', \dots$	Designations for phases in the lithium-hydrogen niobate system that occurs for increasing values of the hydrogen fraction x that occur for x larger than about 0.12.		
β	The phase constant in radian/meter for optical propagation in an optical waveguide. Equal to c/n		
γ	Velocity mismatch scaling parameter $2\pi f \Delta n L$.		
$\Delta\beta$	The difference between the velocities in two parallel optical waveguides (directional coupler) under an applied voltage.		
Δn_e	Change in the extraordinary refractive index of a crystal		

Δn_0	Change in the ordinary refractive index of a crystal
Δz	The length of a computational segment in analysis of velocity mismatch in modulators.
ϵ	Permitivity of a dielectric.
ϵ_0	Permitivity of free space, 8.854×10^{-12} Farad/meter.
ϵ_r	Relative permitivity, ϵ/ϵ_0 ; the dielectric constant.
λ	Wavelength in free space
λ_g	Wavelength in an electromagnetic waveguide; the guide wavelength.
θ	Electrical length, in radians, of a device.

Bandwidth of Linearized Electrooptic Modulators

Uri V. Cummings and William B. Bridges, *Life Fellow, IEEE, Fellow, OSA*

Abstract—Many schemes have been proposed to make high dynamic range analog radio frequency (RF) photonic links by linearizing the transfer function of the link's modulator. This paper studies the degrading effects of finite transit time and optical and electrical velocity dispersion on such linearization schemes. It further demonstrates that much of the lost dynamic range in some modulators may be regained by segmenting and rephasing the RF transmission line.

Index Terms—Bandwidth, electrooptic, linearized, modulators, photonic-link.

I. INTRODUCTION

ELECTROOPTIC intensity modulators have inherently nonlinear transfer functions which may limit the dynamic range of the photonic link through the production of harmonic and intermodulation distortion. Many schemes have been proposed to reduce the distortion byproducts of these modulators by linearizing their transfer functions; for example see the review paper by Bridges and Schaffner [1], and the references therein. The proposed applications for the resulting high dynamic range links include antenna remoting, photonic-coupled phased-array antennas, and cable television transmission.

Linearizing a modulator is a challenge. All of the proposed linearization schemes involve the cancellation of selected distortion terms, and this cancellation depends critically on modulator device parameters. Electrical biases very likely will require active control, and in some modulators, radio frequency (RF) or optical levels will require more accuracy than is realizable with current fabrication techniques.

For applications higher than 1 GHz, traveling wave electrode structures are mandatory to overcome the limitations resulting from interelectrode capacitance and finite transit time. A further difficulty with some popular electrooptic materials, such as lithium niobate, is that the electrical and optical waves travel at different velocities over the finite interaction length of the device, a result of material dispersion. This property limits the modulation-index \times voltage product at high frequencies.

Given the critical dependence of the linearization scheme on modulator parameters, it is a fair question to ask, "How will velocity mismatch effect the linearization results?" This paper addresses that question for several popular modulator types. The summary result is that good velocity matching is essential to successfully linearize some but not all of the

TABLE I
THE PARAMETERS OF A CANONICAL OPTICAL LINK

Velocity Mismatch	Δn	1.8	-
Mod. Length	l_m	10	mm
Laser Power	P_L	100	mW
Laser Noise	R_{IN}	-165	dB/Hz
Optical Loss	L_o	-10	dB
Mod. Sensitivity	V_x, V_s	10	V
Mod. Impedance	R_M	50	Ω
Det. Responsivity	η_D	0.7	A/W
Det. Load	R_D	50	Ω
Noise Bandwidth	BW	1 or 10^6	Hz

modulators. The details differ significantly from one modulator type to another. This paper treats the frequency dependence of six modulator configurations: a standard Mach-Zehnder modulator (MZM), a dual parallel Mach-Zehnder modulator (DPMZM) and a dual series Mach-Zehnder modulator (DSMZM), a simple directional coupler modulator (DCM) at two different bias points, and a directional coupler modulator linearized with two additional dc biased directional couplers in series optically (DCM2P). The results of linearizing these modulators (except for the DSMZM) without regard for transit time were reported in [1]. Now, we report the comparisons including transit time and velocity mismatch.

II. LINK MODEL

This paper assumes the same simplified photonic link and parameters as [1], but now adds the parameters for the finite frequency calculations. They are the effective index mismatch Δn , the modulator length l_m , and the frequency f_0 . The link consists of a laser, an electrooptic modulator, a length of fiber, and a detector. The model excludes electronic preamplification and postamplification. Table I shows all of the parameters associated with this link. These parameters are assumed to have no frequency dependence in the model. The parameters for the active length and the velocity mismatch are typical for LiNbO_3 modulators using simple parallel strip electrodes with no velocity matching. That is, $n_{\text{microwave}} \approx 4.0$ and $n_{\text{optical}} \approx 2.2$, and thus $\Delta n \approx 1.8$. Velocity matching will result in a lower value of Δn .

As in [2], the results are calculated numerically, since no closed-form solution exists for the transfer function of some modulator types. A program, written in C, calculates the frequency-dependent gain and dynamic range. A two-tone electrical test signal, with frequencies f_1 and f_2 drives the modulator. The Fourier transform of the output is evaluated to

Manuscript received October 24, 1997; revised April 21, 1998. This work was supported in part by the United States Air Force Rome Laboratories under Contract F30602-C-96-0020.

The authors are with the California Institute of Technology, Pasadena, CA 91125 USA.

Publisher Item Identifier S 0733-8724(98)05646-1.

find the gain, the harmonic content, and the intermodulation content of the link.

Let $P(p_{in}, t)$ be the electrical signal power after the detector, given the modulator RF drive power p_{in} . Let $\tilde{P}(p_{in}, f)$ be the Fourier transform of $P(p_{in}, t)$.

The gain is

$$\text{Gain}_{\text{dB}} = 10 \{\log[\tilde{P}(p_{in}, f_1)] - \log(p_{in})\}. \quad (1)$$

The small signal gain is obtained by evaluating (1) at sufficiently small p_{in} that the log-log plot of $\tilde{P}(p_{in}, f_1)$ is linear with slope one. In practice, a good value of p_{in} for this calculation is the geometric mean of p_{sat} (the power that drives the modulator voltage to about V_π) and the precision of double precision floating point numbers, or about $p_{in} = -100 \text{ dBm}$.

The spur free dynamic range, DR_{dB} , is the power interval that spans the input power level at which the signal is just distinguishable from the noise and the input power level at which the strongest distortion term becomes distinguishable from the noise. The calculation of DR_{dB} is

$$R_{\text{dB}}(p_{in}) = 10 \max \{\log[\tilde{P}(p_{in}, 2f_1 - f_2)] - \log[\tilde{P}(p_{in}, 2f_1)]\} - \text{noised}_{\text{dB}} \quad (2)$$

$$p_0 = \min(p_{in} | R_{\text{dB}}(p_{in}) = 0) \quad (3)$$

$$\text{DR}_{\text{dB}} = 10 \log[\tilde{P}(p_0, f_1)] - \text{noised}_{\text{dB}}. \quad (4)$$

$R_{\text{dB}}(p_{in})$ is the maximum of the relevant distortion terms minus the noise level in dB. Of the roots of $R_{\text{dB}}(p_{in})$, p_0 is the root that occurs at the lowest power level. DR_{dB} is the difference between p_0 and the input power level at which the signal intersects the noise floor. Since the log-log plot of the signal has slope one, this interval is equivalent to the signal power minus the noise power at the RF drive power at which the distortion power equals the noise level which is (4). Fig. 1 describes the dynamic range calculation for a simple Mach-Zehnder modulator. While dynamic range generally refers to all harmonics and intermodulation products, in practice, there are two dominant distortion terms, the second harmonic, $\tilde{P}(p_{in}, 2f_0)$, and the first intermodulation product, $\tilde{P}(p_{in}, 2f_0 - f_1)$. The linearized modulators in this paper are separated into two categories, each with a different definition of dynamic range. Equation (2) applies to broadband, or *super octave* modulators. That is, $R_{\text{dB}}(p_{in})$ is the maximum of the second harmonic and the intermodulation product. $R_{\text{dB}}(p_{in})$ is defined differently in narrow band or *sub octave* modulators. In this second category, $R_{\text{dB}}(p_{in})$ equals only the third order intermodulation product. In ordinary modulators, the log-log plots of the distortion terms are linear intersecting the noise floor only once. In linearized modulators, the distortion terms are nulled at discrete power levels. They may cross the noise level more than once. It is then necessary to find all of the roots of the distortion minus the noise, R_{dB} , and take the root representing the lowest RF drive power in the definition of dynamic range.

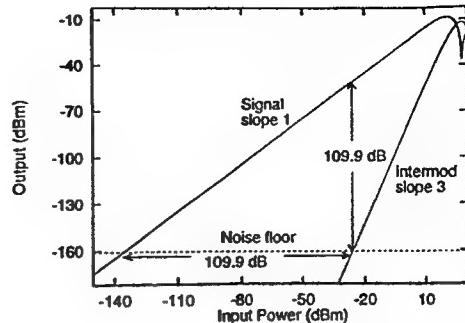


Fig. 1. The signal and third-order intermodulation product of a simple Mach-Zehnder modulator. Since there is effectively no second-harmonic distortion, R_{dB} equals the intermodulation product. R_{dB} crosses the noise floor only once at an input power level of $p_0 = -26 \text{ dBm}$. Since the signal has slope one, the dynamic range may be found from the difference between the signal and p_0 either horizontally or vertically in the plot, giving the familiar dynamic range triangle.

A. Computation of Modulators with Velocity Dispersion

The frequency dependent output of any modulator with a known dc transfer function is calculable. Farwell gives a detailed computational method for this in [2]. The mathematics are straightforward. Let $A_{\text{in}}(t)$ and $A_{\text{out}}(t)$ be the input and output complex amplitudes of the optical wave in a modulator. Let $H(V, l_m)$ be the transfer function, where V is the normalized modulator drive voltage and l_m is the active length of the modulator. If the optical and electrical waves travel at the same velocity, or if the operating frequency is so low that V is effectively constant over l_m , then the output is given by

$$A_{\text{out}}(t) = H[V(t), l_m]A_{\text{in}}(t). \quad (5)$$

Even if there is significant velocity dispersion while the two waves travel the distance l_m , over a short enough section of the guide (Δx), the change in the complex optical amplitudes may be described with the dc transfer function. This is the basis for the frequency dependent calculation. Let x be the coordinate along the optical waveguide, and let t be time. Then

$$\lim_{\Delta x \rightarrow 0} \frac{A(x - \Delta x, t - \Delta t)H[V(x - \Delta x, t - \Delta t), \Delta x]}{A(x, t)} \rightarrow A(x, t). \quad (6)$$

The optical and electrical signals are now functions of two variables, x and t . In (6) the elapsed time equals the incremental length divided by the electrical velocity $\Delta t = \Delta x/v$.

Let there be N equally spaced increments of x and M equally spaced increments of t , and let $x_0 = 0$ and $x_N = l_m$. The modulator is divided into N sections over which the optical and electrical fields are approximately constant. The finite product of the resulting N unitary dc transfer matrices gives the overall transfer function from $A_{\text{in}}(t)$ to $A_{\text{out}}(t)$. Let $A(x_0, t) = A_{\text{in}}(t)$, and $A(x_N, t) = A_{\text{out}}(t)$, and note that $A_{\text{out}}(t) = 0$ for $t < x_N/v$. The approximate modulator

transfer function is

$$A_{\text{out}}(t_j) \approx \left\{ \prod_{k=0}^{N-1} H \left[V \left(x_k, t_j - \frac{x_{N-1-k}}{\nu}, \frac{l_m}{N} \right), \frac{l_m}{N} \right] \right\} A_{\text{in}}(t). \quad (7)$$

The function $V(x, t)$ representing the two tone test is

$$V(x_k, t_j) = p_{\text{in}} \left[\cos \left(\frac{j}{M} - \frac{i}{N} \cdot \frac{\gamma}{f_1} \right) + \cos \left(\frac{j}{M} - \frac{i}{N} \cdot \frac{\gamma}{f_2} \right) \right] \quad (8)$$

$$\gamma \equiv \frac{f_0 \Delta n l_m}{c}. \quad (9)$$

The parameter γ depends on the operating frequency f_0 , the difference between the optical and electrical indices, Δn , the active length l_m , and the velocity of light c . It is important to note that the calculation results depend solely on γ and not on Δn , l_m , and f_0 independently. Thus the results of different lengths or relative wave velocities at different frequencies will be the same if γ is the same. The curves shown below are universal in the sense that they apply to more than the "worst case" velocity mismatch, which is $\Delta n = 1.8$ for lithium niobate modulators. Any change in Δn or l_m leads to a rescaling of the frequency axis for the gain and dynamic range plots.

Equation (7) is general and is the basis for the frequency dependent computations in the model. However, when the active region of a modulator consists of only simple phase shifts, as it does for the Mach-Zehnder modulator, a further simplification may be made. The transfer function is just a diagonal matrix of exponentials. Instead of multiplying exponentials, their arguments are summed. This is equivalent to integrating the location variable out of the voltage function. That is

$$A_{\text{out}}(t) = H \left[\int_0^{l_m} V \left(x, t - \frac{l_m - x}{\nu} \right) dx, l_m \right] A_{\text{in}}(t) \quad (10)$$

$$A_{\text{out}}(t_j) \approx H \left[\sum_{k=0}^{N-1} V \left(x_k, t_j - \frac{x_{N-1-k}}{\nu}, l_m \right), l_m \right] A_{\text{in}}(t). \quad (11)$$

The approximation introduced in (11) comes from the substitution of a summation for an integral, the relationship between A_{out} and H shown in (10) is exact. It may seem odd to use an approximation for a function for which a trivial analytic solution exists [integral of (8)]. However, this is done to mirror the calculation technique for directional couplers and to support the modeling of modulator voltage functions which may not have an analytic integral representation.

In the C-program mentioned above, the temporal increments, M are restricted to be powers of two, so that a radix-2 fast Fourier transform (FFT) algorithm may be used for the spectral analysis.¹ Fig. 2 shows the convergence of the gain of a Mach-Zehnder modulator as a function of spatial increments N at 5, 10, 20, and 40 GHz. The error is normalized, and

¹ It is customary to use the FFT algorithm which is $O(N \log(N))$ instead of the DFT algorithm which is $O(N^2)$. However, it should be noted that the algorithm to compute the modulator output is $O(N^2)$, so the time spent in the FFT algorithm is inconsequential.

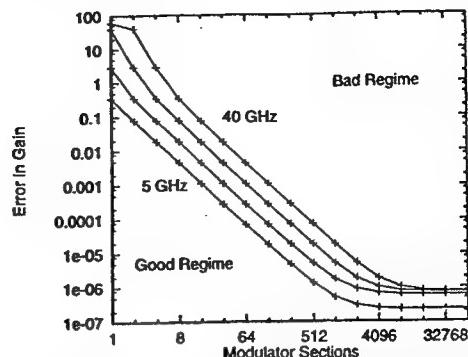


Fig. 2. The convergence of the calculation of the gain of a Mach-Zehnder modulator (MZM), with an increasing number of modulator sections. The error is the magnitude of the calculated gain (not in dB) minus the analytical value normalized by the dc analytical value for the gain. The curves of the log-log plot are for 5, 10, 20, and 40 GHz.

thus the y -axis value "one" corresponds to a 3 dB error and "0.1" corresponds to a 0.4 dB error in the calculated gain. The curves in Fig. 2 are linear until the error is six-to-seven orders of magnitude below the dc value of the gain (since the plot is log-log, constant slopes do not indicate geometric convergence). The convergence saturates at a normalized error of about 10^{-6} because the RF drive power for the small signal gain calculation was arbitrarily chosen to be -100 dBm. Whether in the calculation for gain or dynamic range, the numbers have components that differ by 6–7 orders of magnitude. Since these components occupy the same mantissa, there is a loss in accuracy not recovered by the floating decimal point. Double precision numbers must be used to attain a satisfactory accuracy. It is interesting to note that the calculated points form a horizontal line across the curves in the linear regime. This indicates that doubling γ (by doubling the frequency for instance) exactly requires a doubling of the modulator sections to achieve the same error. Efficient code allows the calculation hundreds of frequency points with a 128-point FFT and a comparable number of spatial increments in seconds on a contemporary desk-top machine (120 MHz Pentium processor).

III. MACH-ZEHNDER AND DIRECTIONAL COUPLER MODULATORS

The most common electrooptic modulator is the Mach-Zehnder interferometer (MZM). It has a sine-squared transfer function, and the gain is a sinc function of the frequency-length-index product, γ . When biased at the half-wave voltage, $V = 0.5V_\pi$, it attains its maximum linearity and dynamic range. All even-order harmonics are identically zero. The intermodulation distortion product solely determines the dynamic range, even in a super-octave system. The dynamic range is independent of frequency; the signal decays with frequency, but the intermodulation product decays identically. Thus, the range of RF drive powers (in dB), over which there are no spurs above the noise, shifts with a change in frequency, but it does not expand or

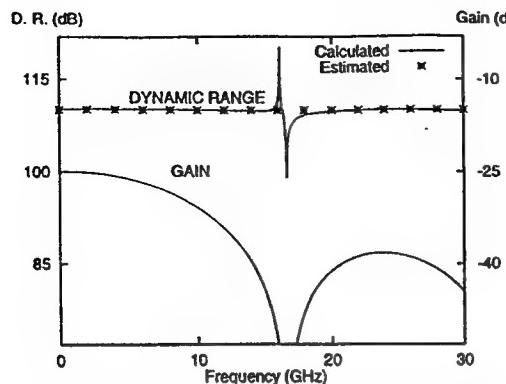


Fig. 3. Gain and dynamic range of a standard Mach-Zehnder modulator. The pole and zero in the dynamic range are the frequencies at which the signal and intermodulation products go to zero, respectively.

contract. Given its analytical simplicity and widespread use, the Mach-Zehnder is used first to evaluate the accuracy of the numerical calculation, and then it is used for a comparison to the linearized modulators.

Fig. 3 shows the calculations for gain and dynamic range of a simple MZM. The gain has the form $[\sin(\pi\gamma)/(\pi\gamma)]^2$ with zeros at multiples of 16.2 GHz (where $\gamma = 1$ for the canonical parameters from Table I), and a low frequency link gain of -25.5 dB (also appropriate for the link parameters). The dynamic range is flat except for a null and singularity near the gain null. This is a simple numerical artifact, resulting from the finite frequency difference between the two tones in the driving function. The signal, f_1 , and the intermodulation product, $2f_1 - f_2$, are at slightly different frequencies, and hence, they null at different frequencies. The dynamic range goes to zero at the signal null and it goes to infinity at the distortion null.

Fig. 4 shows the analogous calculations for a directional coupler modulator. The low frequency gain is -24.8 dB, 0.7 dB better than Mach-Zehnder.² The first null of the gain of the directional coupler occurs at 26 GHz, $\gamma = 1.6$, compared to 16 GHz for the MZM. The first lobe of the gain curve does not correspond to the sinc function of the Mach-Zehnder. However, subsequent nulls are periodic with a 16 GHz period, resulting from the underlying γ of the directional coupler. The frequency at which the gain has fallen by 3 dB is 40% higher than that of the Mach-Zehnder modulator with the same index-length product.

The dynamic range compares unfavorably to that of the Mach-Zehnder. At low frequency it is similar to that of the MZM, and it is approximately flat with frequency. However there is a kink in the curve at 1.8 GHz (left vertical arrow), after which the dynamic range decays rapidly with frequency. Unlike the Mach-Zehnder, where all even-order derivatives of the transfer function are identically zero when the modulator

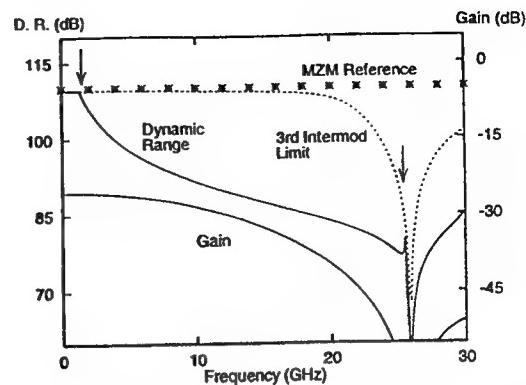


Fig. 4. Gain and dynamic range of a directional coupler modulator biased at the $0.43 V_r$ point. The two arrows indicate that the dynamic range changes from intermodulation-limited to second-harmonic limited or vice versa. If the second harmonic is ignored (suboctave operation), the dynamic range follows the dashed curve.

is biased at $0.5V_r$, in the directional coupler only the second derivative is zero at its optimal bias, $0.43V_s$. Distortion migrates from nonzero fourth-, sixth-, eighth-, etc., order derivatives to the frequency at which the second harmonic occurs. It does so rapidly with increasing frequency, and it equals the third order intermodulation product at 1.8 GHz. At higher frequencies, distortion from these nonzero even derivatives limits the dynamic range. The intermodulation-limited dynamic range (as if this were a suboctave modulator) is shown with a dotted line to demonstrate the intermingling of the second harmonic and third order intermodulation product. The bias voltage that nulls the total second harmonic, which arises from all even order derivatives, is thus a function of frequency in the DCM, but it is independent of frequency in the MZM. At any finite frequency, there is a bias value which nulls the total second harmonic, but the null will hold only over a narrow bandwidth. While the gain compares favorably to the MZM, the dynamic range makes the directional coupler inferior to the Mach-Zehnder for superoctave, high-frequency applications.

IV. BROAD-BAND LINEARIZED MODULATORS

Dynamic range values are calculated for two linearized broad-band, or superoctave, modulators: the dual parallel Mach-Zehnder (DPMZM) and the directional coupler with two passive sections (DCM2P) which are described in [4] and [5], respectively. The DPMZM has two identical, single Mach-Zehnder modulators in parallel optically and electrically but with unequal levels of optical and RF power driving the two modulators. Both modulators are biased at the $0.5V_r$ point, but with opposite slopes, so the modulators are 180 degrees out of phase. Most of the optical power and a small fraction of the RF power drive one modulator. A small amount of optical power and the majority of the RF power drive the other modulator, creating relatively larger distortion products than in the first modulator. The two signals are combined incoherently in the photodetector. The RF and optical splits are adjusted so that the distortion terms cancel exactly, but the

²The comparison of the gain bandwidth product between the MZM and the DCM assumes that $V_r = V_s$; that is, the modulators have the same normalization voltages. While these voltages should be similar in the same manufacturing process, it is hard to directly compare them since the electrode geometries and crystal orientations of the two modulator types may differ.

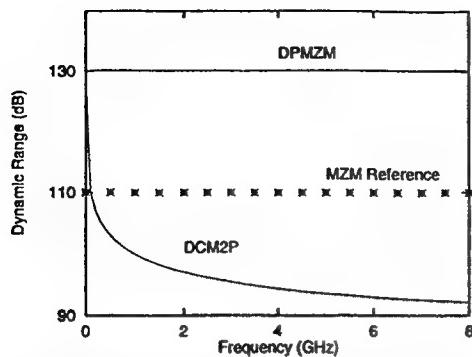


Fig. 5. Dynamic range comparison of two linearized modulators: a dual parallel Mach-Zehnder modulator (DPMZM) and a directional coupler modulator with two passive sections (DCM2P). As in Fig. 3 for the simple MZM the dynamic range of the DPMZM is independent of frequency. However, the dynamic range of the DCM2P drops precipitously with frequency.

signals do not. Since these two paths are in parallel, the effects of velocity mismatch apply equally to each Mach-Zehnder. Thus, the distortion terms still precisely cancel regardless of frequency, and this is the result shown in Fig. 5. While the DPMZM is robust to velocity mismatch, in practice it is hard to make broadband. A precise RF split must be maintained over the desired frequency band. If it varies, the dynamic range will decrease at all but the narrow frequency at which the RF split is optimized.

The linearized directional coupler with two passive sections can be adjusted to provide high dynamic range with both intermodulation and second harmonic reduction in the absence of velocity mismatch. However, it suffers severely from velocity mismatch as shown in Figs. 5 and 6. By 80 MHz it is no better than the ordinary Mach-Zehnder and by 8 GHz it is 18 dB worse than the ordinary MZM. Fig. 6 shows the first 50 MHz of the DCM2P dynamic range in more detail. Unlike the DPMZM, in the DCM2P the mismatch between the RF drive and the modulated signal upsets the critical distortion cancellation conditions.

It has previously been reported [1] that the distortion cancellation condition is critically sensitive to the modulator parameters, particularly bias voltage. The voltage on the bias electrodes must be maintained to a very high accuracy. The accuracy required depends on the operating bandwidth, also explained in [1]. This requires active bias stabilization. Given the critical bias conditions and the fact that the distortion cancellation sections are in series, unlike the DPMZM, the rapid degradation with dynamic range is reasonable. The original experiments on this modulator were performed at audio frequencies where these effects would not be noticed [5]. Subsequently, measurements at 1 and 2 GHz [6] were single-frequency measurements, with the bias values reoptimized for the operating frequency; no bandwidth measurements around 1 and 2 GHz were made.

V. SUBOCTAVE LINEARIZED MODULATORS

For suboctave applications, the second harmonic may be ignored; the third-order intermodulation product alone de-

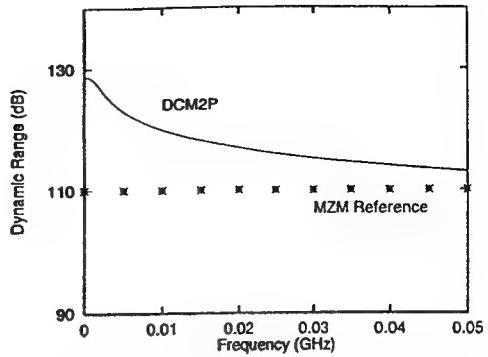


Fig. 6. Expanded plot of the dynamic range of the DCM2P shown in Fig. 5.

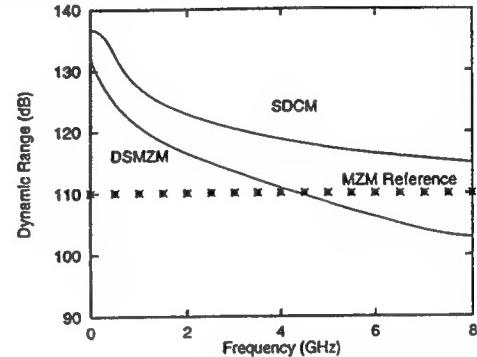


Fig. 7. Dynamic range comparison of two suboctave linearized modulators: a dual series Mach-Zehnder modulator (DSMZM) and a directional coupler modulator biased at $0.79V_s$ (SDCM).

termines the dynamic range. Two suboctave modulators are analyzed in this section.

- 1) The dual series Mach-Zehnder modulator (DSMZM), as described in [7], which has two MZM's in series optically, the same bias voltage on each pair of electrodes, and a single RF electrode covering both modulators.³
- 2) A suboctave directional coupler modulator (SDCM), i.e., a directional coupler biased at $0.79V_s$ instead of $0.43V_s$, as described in [1]. Unlike the MZM, in the directional coupler, the intermodulation product nulls at a different voltage ($0.79V_s$) than the signal ($1.0V_s$). Thus no extra electrode sections (as in the DCM2P) are needed to make a suboctave directional coupler.

Fig. 7 shows the dynamic range as a function of frequency for the DSMZM and SDCM compared to the standard MZM reference (horizontal dotted line). Both of these modulators suffer from the effects of velocity mismatch and transit time. However, unlike the DCM2P, the DSMZM shows an advan-

³There are other cascaded Mach-Zehnders proposed in the literature. In some there are different bias voltages on the two electrodes. In [8] there is a time delay between the first and second modulator so that the RF drive and the modulated signal are rephased at the second Mach-Zehnder. While this version may be more common in the literature, for the purposes of a fair comparison, this modulation scheme is addressed in the section on periodic rephasing below. Additionally, in [9] a mixed directional coupler and Mach-Zehnder scheme purports to minimize the second and third harmonic.

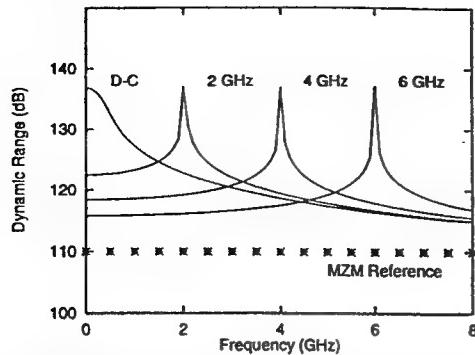


Fig. 8. Reoptimization of the suboctave linearized directional coupler modulator (SDCM). Slightly adjusting the bias voltage around $0.79 V_s$ fully recovers the low-frequency dynamic range optimum, but only over a narrow bandwidth.

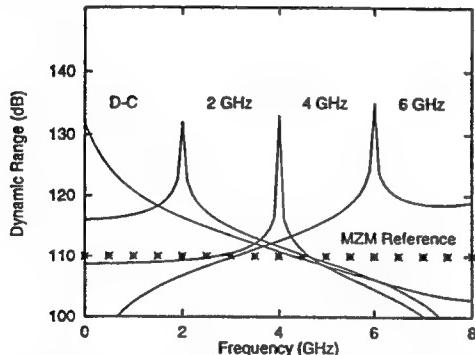


Fig. 9. Reoptimization of the dual series Mach-Zehnder modulator (DSMZM). Adjusting the bias voltage fully recovers the low-frequency dynamic range optimum, but only over a narrow bandwidth.

tage up to 4 GHz over the standard Mach-Zehnder, and the SDCM is better still. The SDCM is consistently about 5 dB better than the DSMZM. At the SDCM bias ($0.79 V_s$), the dc optical output is quite small, thus reducing the shot noise at the detector, which is the dominant term in the total noise (as described in [1]).

The bias voltage of the SDCM controls the frequency of the dynamic range optimum. Small adjustments in the bias around the dc value of $0.79 V_s$ allow the full recovery of the low-frequency dynamic range optimum at any center frequency. Unfortunately, this results in a relatively narrow operating bandwidth around the reoptimization frequency, as shown in Figs. 8 and 9. The critical control of bias voltage will doubtless require the use of pilot tones to minimize the harmonic or intermodulation products. However, the results of Figs. 8 and 9 indicate that these pilot tones will have to be within the band of interest, not at low frequencies.

VI. THE EFFECTS OF NOISE

Noise effects linearized modulators somewhat differently than standard modulators. In a standard MZM, the noise bandwidth reduces the dynamic range by $(BW)^{2/3}$. In Fig. 1 the signal is a line with slope one, and the third-order intermodu-

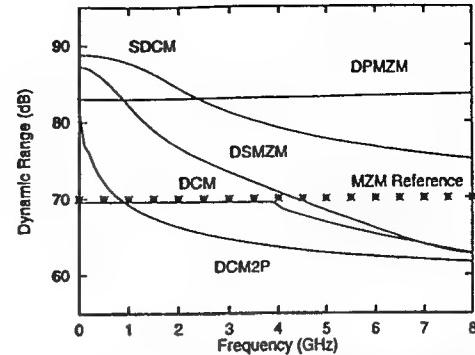


Fig. 10. Dynamic range comparison with 1 MHz noise bandwidth of five modulator configuration: the superoctave dual parallel Mach-Zehnder modulator (DPMZM), the superoctave directional coupler modulator linearized by two passive couplers in series (DCM2P), the suboctave dual series Mach-Zehnder (DSMZM), the suboctave linearized directional coupler (SDCM), and the standard directional coupler modulator (DCM).

lation product is a line with slope three. The "noise floor" is a third, horizontal line which forms a triangle with the signal and the intermodulation line. The length of the base of this triangle is the dynamic range (in dB). The vertical position of the noise line is proportional to the $\log(BW)$, so from simple geometry, it is clear that the dynamic range goes as $(BW)^{2/3}$. However, in a linearized modulator the third-order intermodulation product at $(2f_2 - f_1)$ has been nulled. The dominant intermodulation term is at $(3f_2 - 2f_1)$, and it grows as the fifth power of the RF drive. Thus the noise bandwidth reduces the dynamic range by $(BW)^{4/5}$ (More complicated linearization schemes can result in even steeper slopes for intermodulation, as discussed in [1]). The excess dynamic range, X , of a linearized modulator over a ordinary modulator is

$$X = C - \frac{4}{3} \log(BW) \quad (12)$$

where BW is a the bandwidth in Hertz, and C is a constant in dB equal to the difference between the dynamic range of a linearized modulator and a standard modulator with a 1-Hz noise bandwidth. This equation is only valid in the approximation that the signal and intermodulation curves are straight lines. For a comparison of the suboctave DCM versus the standard MZM, C equals 26 dB. For a 1 MHz noise bandwidth, $X = 18$ dB. The dynamic range is often given for a 1-Hz noise bandwidth, and this simple scaling rule is applied to find the dynamic range of a system with a realistic bandwidth.

Apart from the different scaling rules, the noise bandwidth has an additional effect on linearized modulators. The dynamic range becomes less sensitive to device variations, either bias voltages or fabrication parameters, as the noise bandwidth increases. This makes it easier to maintain the distortion cancellation than is implied by some of the previous figures. Fig. 10 shows the dynamic range for the four modulators discussed above, but now with a 1-MHz noise bandwidth. The DPMZM and the MZM reference lines are flat as in the 1-Hz bandwidth case. The simple DCM is flat out to the

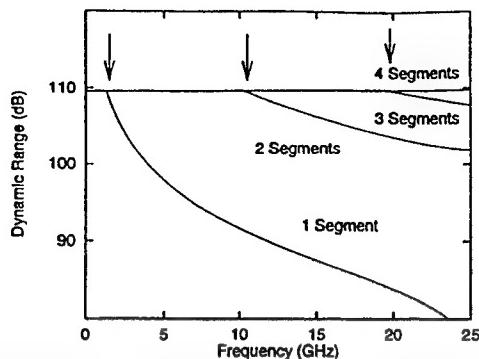


Fig. 11. Dynamic range versus frequency for the standard directional coupler modulator (DCM) with multiple electrode segments. The arrows show the breaks between the intermodulation and the second-harmonic limitation of the dynamic range.

frequency where the second harmonic exceeds the intermodulation distortion, as in Fig. 3. However, the crossover now occurs at 3.8 GHz instead of 1.8 GHz. The DSMZM, SDCM, and DCM2P dynamic ranges, which all roll off with frequency, now do so at a slower rate than in the 1-Hz bandwidth case. While the DCM2P still rolls off too quickly to be a useful modulator, the suboctave modulators are starting to show reasonable frequency performance. A similar reduction in peak dynamic range, with a broadening of the bandwidth over which it occurs, is also obtained in the bias reoptimized curves of Figs. 8 and 9.

VII. BANDWIDTH RECOVERY THROUGH PERIODIC REPHASING

One method of overcoming the degradation in dynamic range due to velocity mismatch is to break the transmission line into a number of segments and rephase the signal at the beginning of each segment. This is velocity-matching "on the average." The technique has been used successfully in a number of forms, for example, [10]–[12]. The program written for this study is easily modified to make calculations for such periodically rephased modulators, since the modulator is already broken up into a cascade of matrices. Thus, the modulator is incrementally velocity mismatched for a few matrices and is then rephased for the next section, and so on. Fig. 11 shows the results of such a calculation for the simple DCM link with the parameters used above, but with the modulator's transmission line having 1, 2, 3, and 4 segments. The 1-segment curve repeats the result in Fig. 4 (no rephasing) for reference. With only two segments (one rephasing), the bandwidth over which the dynamic range is flat improves vastly, and using four segments gives an essentially flat dynamic range. Of course, it would still be preferable to use a standard MZM for broadband links since it does not require rephasing.

A similar dramatic improvement is obtained in the SDCM, biased at the $0.79 V_s$ point, as shown in Fig. 12. The curve from Fig. 7 is shown for reference along with curves for two-, three-, and four-electrode segments. The curve for one segment initially shows a deep roll-off in dynamic range and then a more gradual roll-off, with only 5 dB of dynamic range improvement over the standard MZM remaining at 8 GHz.

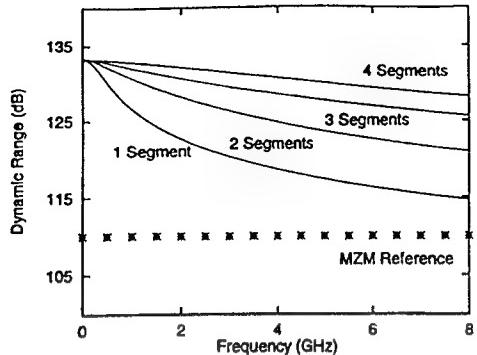


Fig. 12. Dynamic range versus frequency for the suboctave linearized directional coupler (SDCM) with multiple electrode segments.

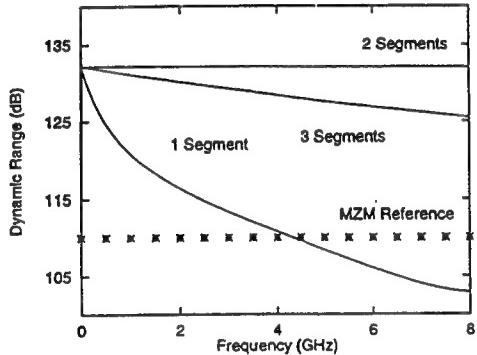


Fig. 13. Dynamic range versus frequency for the suboctave dual series Mach-Zehnder (DSMZM) with multiple electrode segments. Note that two (and any even number) segment give frequency independent dynamic range.

With just two segments, the roll-off is made gradual over the whole range. With four segments, there is 18 dB of dynamic range improvement remaining at 8 GHz. This figure assumes a 1-Hz noise bandwidth. When a 1-MHz noise bandwidth is used, the roll-off is more gradual. For instance, the dynamic range of the two-electrode segment SDCM is better than the two-electrode DSMZM (shown in Fig. 13) up to 5.9 GHz.

A very interesting result of the application of rephasing is seen in the dual series Mach-Zehnder modulator. This modulator has been shown to give very high dynamic range values based on the intermodulation distortion, although it has, like the SDCM, a very large second harmonic content, restricting it to suboctave applications as stated in [8]. The decrease in dynamic range with frequency is shown in Fig. 13 in the curve labeled "1 segment" meaning one set of traveling wave electrodes spanning two Mach-Zehnder modulators (with an admittedly unrealistic zero distance assumed between the two modulators). This is simply a repeat of the curve in Fig. 7. The surprise is that with only one rephasing (two segments) the modulator exhibits frequency-independent performance. Breaking the electrodes into three segments yields an improvement over one segment, but not as good as two or any even number of segments, all of which yield perfect performance. This behavior results from the symmetry of the modulators and their bias points.

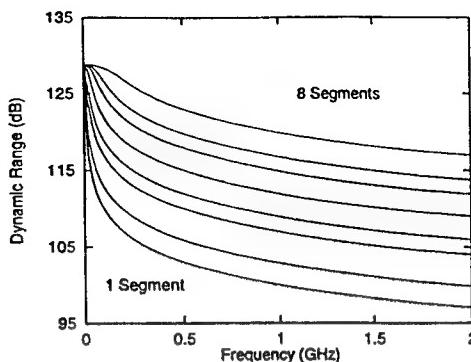


Fig. 14. Dynamic range versus frequency for the superoctave linearized directional coupler (DCM2P) with one to eight electrode segments. While rephasing yields improvements, the results remain inferior to other modulators.

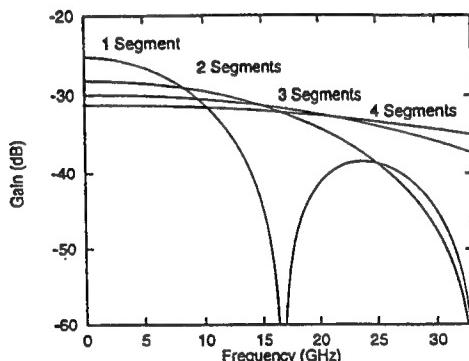


Fig. 15. The gain of a standard Mach-Zehnder with multiple electrode segments.

The results for the DCM2P superoctave modulator remain discouraging, even with rephasing. Fig. 14 shows the dynamic range from one-to-eight repated electrode segments. This plot spans 2 GHz only, which is enough to capture the interesting range. Even with eight electrodes, the frequency at which the dynamic range improvement over the simple MZM has fallen 10 dB is only 1.3 GHz. While this plot assumes an unrealistically low 1 Hz noise bandwidth, it also assumes perfect bias control. Given the modulators sensitivity to its three bias voltages, it is unlikely that this modulator is suitable for microwave applications with reasonable bandwidth.

Segmenting the RF electrode does have a low-frequency gain penalty. While flattening the gain versus bandwidth curve, it also lowers the absolute gain at low frequencies. The reduction occurs from the splitting of the RF power. It is split N ways to feed the N electrodes, reducing the voltage by \sqrt{N} on each electrode segment. However, if the rephasing is achieved by a series intermittent interaction method such as described in [12], then there is no power splitting and subsequently no $1/\sqrt{N}$ penalty. But this scheme involves long, curved, transmission lines and thus may have unacceptable losses. Fig. 15 shows the calculated gain of a simple Mach-Zehnder with one-, two-, three-, and four-electrode segments. Above some cross-over frequency each curve has a better absolute gain than the curve with one fewer electrode segments. Clearly, true

velocity matching, by somehow making $\Delta n \rightarrow 0$ is preferable if possible, since there is no \sqrt{N} penalty.

We wish to remind the reader that all the results obtained in this study may be applied to modulators with any degree of velocity matching by rescaling the frequency axis by the change in γ , the frequency-length-index product.

ACKNOWLEDGMENT

The authors would like to thank E. Ackerman, G. E. Betts, and C. H. Cox III of Lincoln Laboratories, Cambridge, MA, N. P. Bernstein of the Air Force Rome Laboratory, and J. H. Schaffner of Hughes Research Laboratories. The authors also wish to acknowledge the support and enthusiasm of the late B. Hendrickson of Rome Laboratories and ARPA.

REFERENCES

- [1] W. B. Bridges and J. H. Schaffner, "Distortion in linearized electrooptic modulators," *IEEE Trans. Microwave Theory and Tech.*, vol. 43, pp. 2184-2197, Sept. 1995.
- [2] M. L. Farwell and W. S. C. Chang, "Simulating the response of coupled channel and interferometric modulator designs," *J. Lightwave Technol.*, vol. 13, Oct. 1995.
- [3] B. Pucel, L. Riviere, and J. L. Bihani, "New model for the active directional coupler," *J. Lightwave Technol.*, vol. 14, pp. 1501-1506, June 1996.
- [4] S. K. Korotky and R. M. Ridder, "Dual parallel modulation schemes for low-distortion analog optical transmission," *IEEE J. Select. Areas Commun.*, vol. 8, pp. 1377-1381, Sept. 1990.
- [5] M. L. Farwell, Z. Q. Lin, E. Wooten, and W. S. C. Chang, "An electrooptic intensity modulator with improved linearity," *IEEE Photon. Technol. Lett.*, vol. 3, pp. 792-795, Sept. 1991.
- [6] M. L. Farwell, W. S. C. Chang, and J. H. Schaffner, "A directional coupler modulator with improved linearity (operating up to 1 GHz)," in *Leos Summer Topical Meet. Dig. Broadband Analog and Digital Optoelectron.*, 1992, vol. WA3, pp. 8-9.
- [7] G. E. Betts and F. J. O'Donnell, "Microwave analog optical links using suboctave linearized modulators," *IEEE Photon. Technol. Lett.*, vol. 8, pp. 1273-1275, Sept. 1996.
- [8] G. E. Betts, "Linearized modulator for suboctave-bandpass optical analog links," *IEEE Trans. Microwave Theory Tech.*, vol. 42, pp. 2642-2649, Dec. 1994.
- [9] H. Skeie and R. V. Johnson, "Linearization of electrooptic modulators by a cascade coupling of phase modulating electrodes," in *Proc. SPIE*, 1991, vol. 1583, pp. 153-164.
- [10] W. B. Bridges, F. T. Sheehy, and J. H. Schaffner, "Wave-coupled electrooptic modulators for microwave and millimeter-wave modulation," *IEEE Photon. Technol. Lett.*, vol. 3, pp. 133-135, Feb. 1991.
- [11] R. C. Alfemess, S. K. Korotky, and E. A. J. Marcatili, "Velocity-matching techniques for integrated optic traveling wave switch modulators," *IEEE J. Quantum Electron.*, vol. QE-20, pp. 301-309, Mar. 1984.
- [12] J. H. Schaffner, "Analysis of a millimeter wave integrated electro-optic modulator with a periodic electrode," in *Proc. SPIE OE-LASE Conf.*, 1990, vol. 1217, Los Angeles, CA, Jan. 16-17, 1990, pp. 101-110.



Uri V. Cummings was born in Vallejo, CA, on November 9, 1971. He received the B.A. degree in English from Wesleyan University, Middletown, CT, and the B.S. degree from the California Institute of Technology (Caltech), Pasadena, simultaneously in 1994. He received the M.S. degree in electrical engineering from the Caltech in 1995. Currently, he is working toward the Ph.D. degree in electrical engineering at Caltech.

His research interests include the study of linearized electrooptic modulators and the design and experimental verification of antenna-arrayed electrooptic modulators at 100 GHz.



William B. Bridges (S'53-M'61-F'70-LF'98) was born in Inglewood, CA, in 1934. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1956, 1957, and 1962, respectively. His graduate research dealt with noise in microwave tubes and electron-stream instabilities (which later became the basis of the Vircator.)

From 1957 to 1959, he was an Associate Professor in the Department of Electrical Engineering at the University of California, Berkeley, teaching courses in communication and circuits. Summer jobs at RCA and Varian provided stimulating experience with microwave radar systems, ammonia beam masers, and the early development of the ion vacuum pump. In 1960, he joined the Hughes Research Laboratories, Malibu, CA, as a Member of the Technical Staff, and from 1968 to 1977, he was a Senior Scientist with a brief tour as Manager of the Laser Department in 1969-1970. His research at Hughes Research Laboratories involved gas lasers of all types and their application to optical communication, radar, and imaging systems. He is the discoverer of laser oscillation in noble gas ions and spent several years on the engineering development of practical high-power visible and ultraviolet ion lasers for military applications. He joined the faculty of the California Institute of Technology, Pasadena, CA, in 1977 as Professor of Electrical Engineering and Applied Physics, and serving as Executive Officer for Electrical Engineering from 1979 to 1981. In 1983, he was appointed Carl F. Braun Professor of Engineering and conducted research in optical and millimeter-wave devices and their applications. Current studies include the millimeter-wave modulation of light and high-fidelity analog microwave photonic links. He is coauthor (with C. K. Birdsall) of *Electron Dynamics of Diode Regions* (New York: Academic, 1966).

Dr. Bridges is a member of Eta Kappa Nu, Tau Beta Pi, Phi Beta Kappa, and Sigma Xi, receiving Honorable Mention from Eta Kappa Nu as an "Outstanding Young Electrical Engineer" in 1966. He received the Distinguished Teaching Award in 1980 and 1982 from the Associated Students of Caltech, the Arthur L. Schawlow Medal from the Laser Institute of America in 1986, and the IEEE LEOS Quantum Electronics Award in 1988. He is a member of the National Academy of Engineering and the National Academy of Sciences, and a Fellow of the Optical Society of America (OSA) and the Laser Institute of America. He was a Sherman Fairchild Distinguished Scholar at Caltech, Pasadena, CA, in 1974-1975, and a Visiting Professor at Chalmers Technical University, Göteborg, Sweden, in 1989. He has served on various committees of both IEEE and OSA and was formerly Associate Editor of the IEEE JOURNAL OF QUANTUM ELECTRONICS and the *Journal of the Optical Society of America*. He was the President of the Optical Society of America in 1988, a member of the United States Air Force Scientific Advisory Board 1985-1989, and a member of the Board of Directors of Uniphase Corporation 1986-1998.



US005886807A

United States Patent [19]
Cummings

[11] Patent Number: 5,886,807
[45] Date of Patent: Mar. 23, 1999

[54] TRAVELING-WAVE REFLECTIVE
ELECTRO-OPTIC MODULATOR

3,791,718 2/1974 Chow 359/261
5,076,655 12/1991 Bridges 385/3
5,291,565 3/1994 Schaffner et al. 385/3
5,355,422 10/1994 Sullivan et al. 385/2
5,548,668 8/1996 Schaffner 385/2
5,790,719 8/1998 Mitomi et al. 385/2

[75] Inventor: Uri Cummings, Pasadena, Calif.

[73] Assignee: California Institute of Technology,
Pasadena, Calif.

[21] Appl. No.: 999,194

FOREIGN PATENT DOCUMENTS

63-141021 6/1988 Japan 385/2
01-232323 9/1989 Japan 385/3
03-262012 9/1991 Japan 385/9
05-249419 9/1993 Japan 385/2
09-015545 1/1997 Japan 385/3

Related U.S. Application Data

[60] Provisional application No. 60/036,243 Jan. 24, 1997.

Primary Examiner—Georgia Epps

[51] Int. Cl.⁶ G02F 1/35

Assistant Examiner—Margaret Burke

[52] U.S. Cl. 359/263; 359/245; 359/251;
359/254; 385/2; 385/3; 385/8; 385/9

Attorney, Agent, or Firm—Fish & Richardson P.C.

[58] Field of Search 359/245, 251,
359/254, 255, 259, 261, 263; 385/2, 3,
8, 9

[57] ABSTRACT

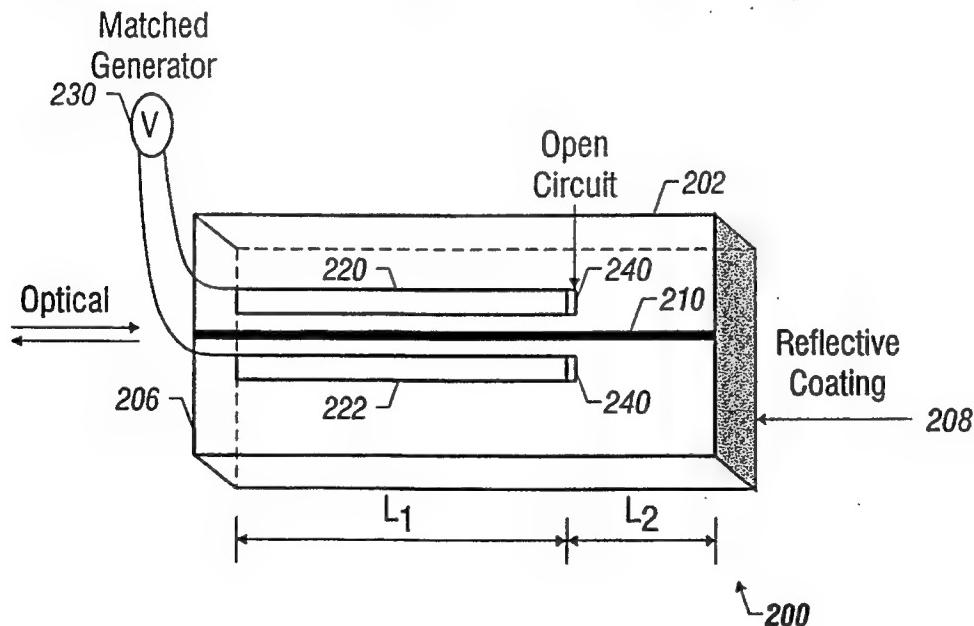
[56] References Cited

A traveling-wave electro-optic modulator using reflected signals to achieve a one-time rephasing. For a given crystal length, the interaction length is effectively doubled and a 3-dB gain is automatically achieved compared to a conventional traveling-wave modulator.

U.S. PATENT DOCUMENTS

3,694,052 9/1972 Ortel 359/261

29 Claims, 4 Drawing Sheets



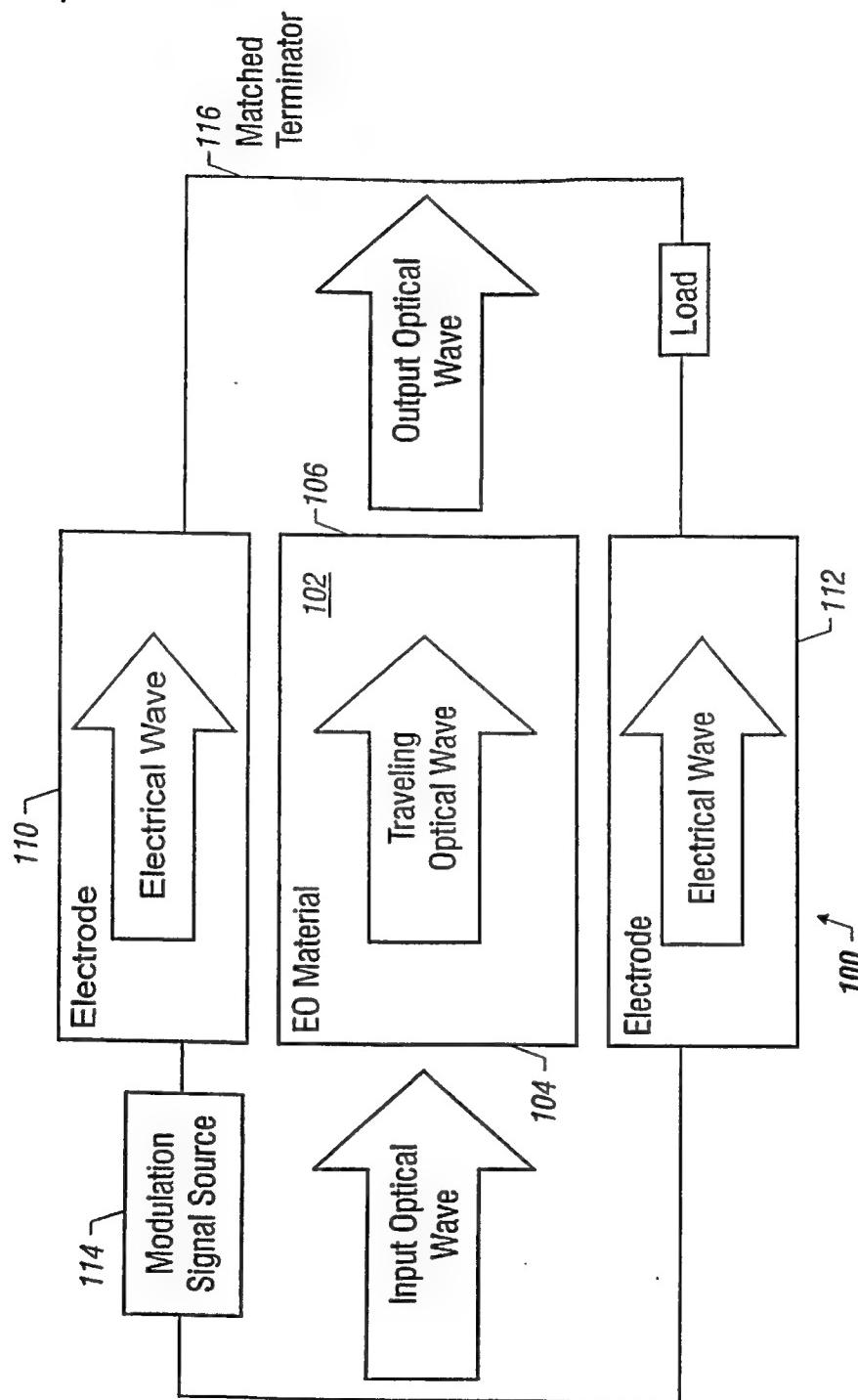


FIG. 1
(PRIOR ART)

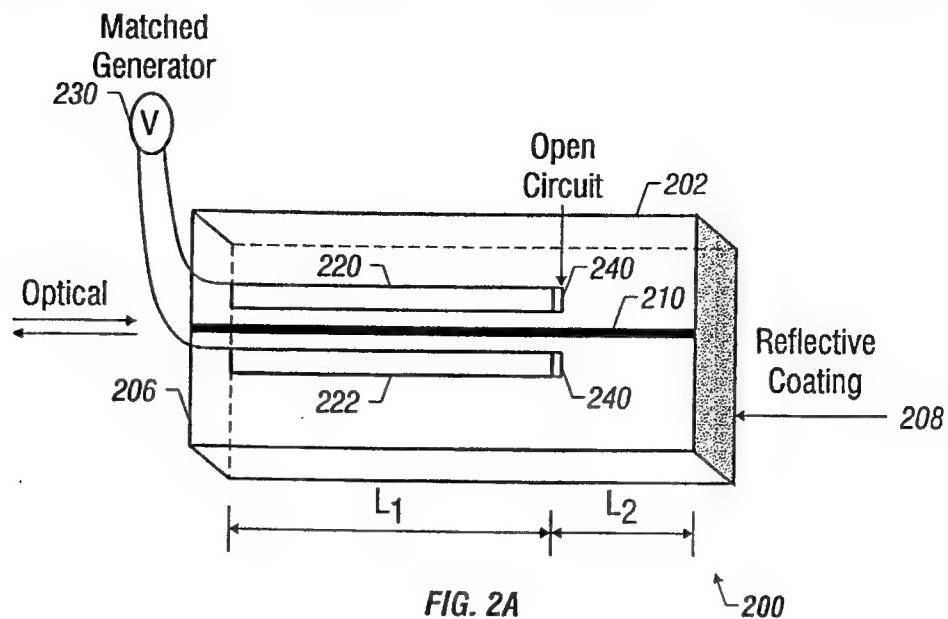


FIG. 2A

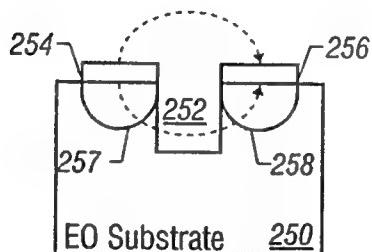


FIG. 2B

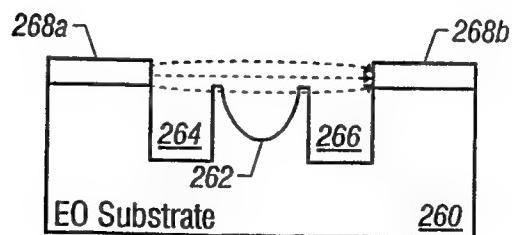
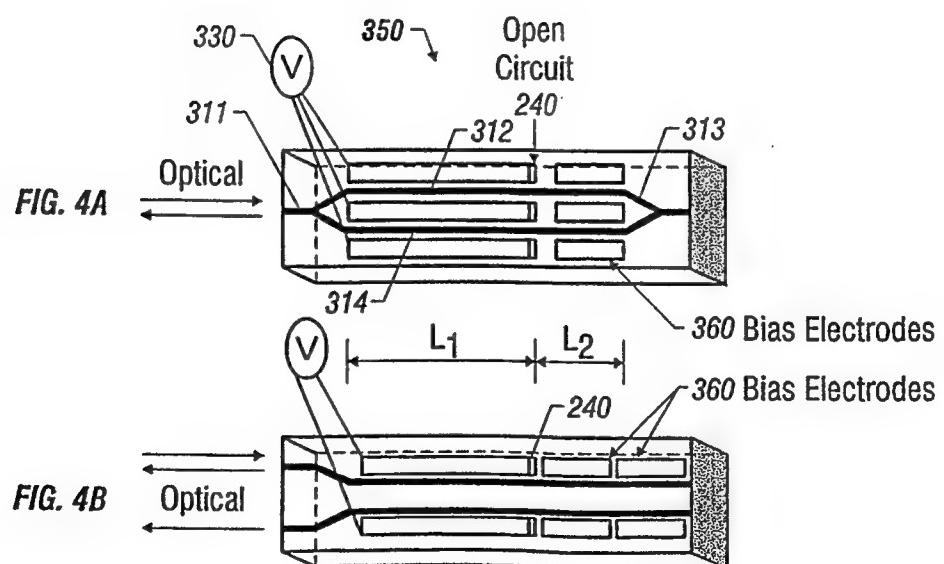
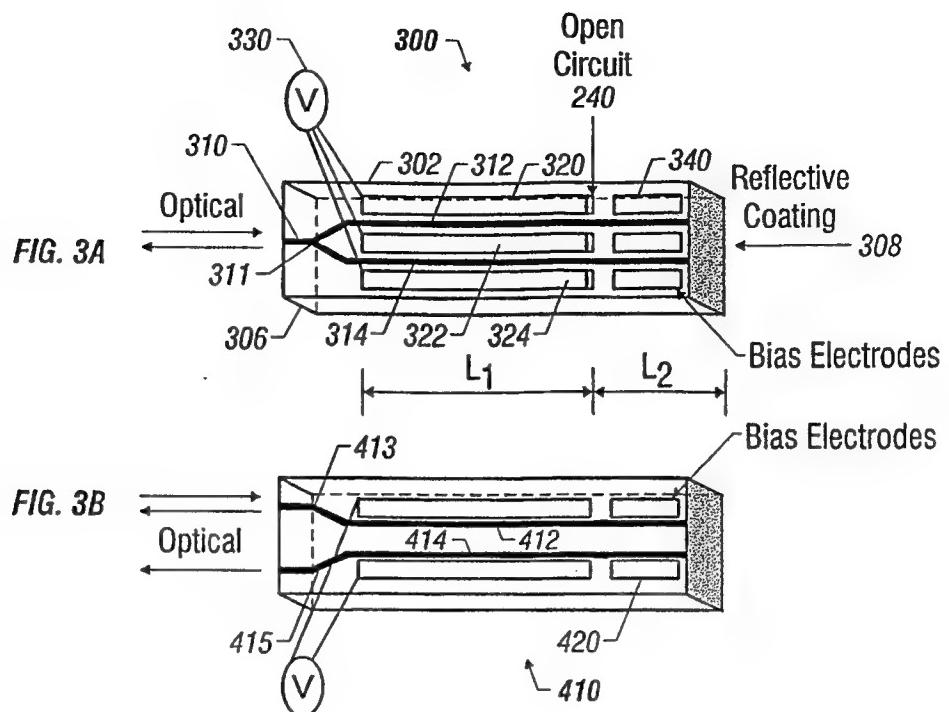


FIG. 2C



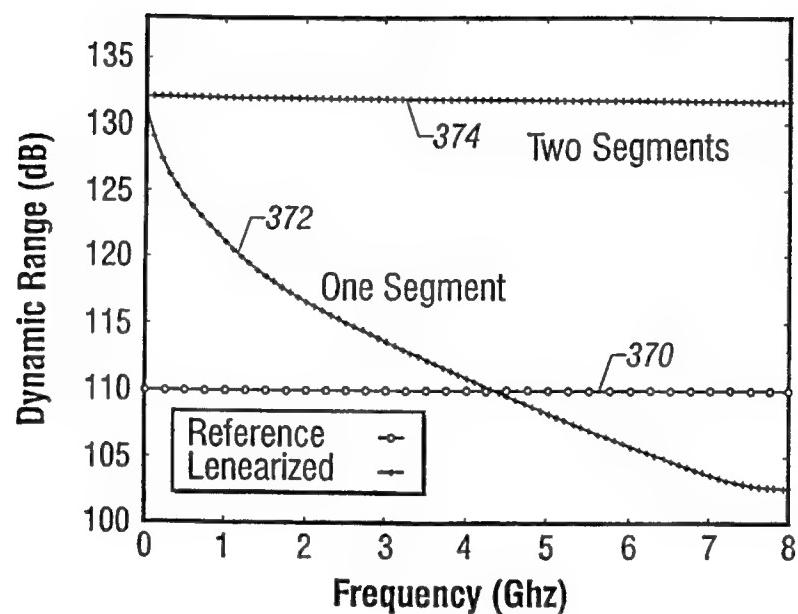


FIG. 3C

I

- TRAVELING-WAVE REFLECTIVE
ELECTRO-OPTIC MODULATOR

ORIGIN OF THE INVENTION

The U.S. Government has certain rights to this invention pursuant to Grant No. F30602-C-96-0020 awarded by the United States Air Force.

This application claims the benefit of U.S. Provisional Application No. 60/036,243 filed on Jan. 24, 1997, which is incorporated herein by reference in its entirety.

FIELD OF THE INVENTION

The present invention relates to electro-optic modulators, and more specifically, to a traveling-wave electro-optic modulator for minimizing phase mismatch between the electrical and optical waves within such a modulator.

BACKGROUND OF THE INVENTION

Light modulation can be programmably controlled by electrical signals and has been widely used in data processing devices and communication systems (e.g., cable TV). One implementation of programmable light modulation is electro-optic ("EO") light modulation by using an electro-optic material (e.g., crystals) whose index of refraction changes in response to an externally applied electrical field.

Traveling-wave EO modulators are suitable for light modulation at high frequencies such as in a range above $\sim 10^8$ Hz. Light modulation using traveling-wave modulators is known to the art. See, for example, "Optical Waves in Crystals: propagation and Control of Laser Radiation", by Yariv and Yeh, John Wiley and Sons (1980). In a traveling-wave EO modulator, the direction of the electrical field of a modulating field is usually perpendicular to the propagation direction of the optical wave. In particular, the modulating wave travels along with the optical wave in substantially the same propagation direction. Much effort in research and development has been made in designing traveling-wave modulators due to their applications in data processing and communication devices and systems. Compact integrated EO modulators are of particular interest in this field for optical fiber networks and other applications.

FIG. 1 shows a conventional traveling-wave light modulator 100. An EO material 102 of length L with optical input and output facets 104 and 106 is placed between two parallel electrodes 110 and 112. In an integrated modulator configuration, the EO material 102 forms a substrate on which an optical waveguide and the electrodes are formed. A modulation signal source 114 is connected to the electrodes 110 and 112 at their ends near the input facet 104 to launch a traveling modulation wave at a modulation frequency ω_{mod} . A signal terminator 116 with a load can be used to terminate the modulating waves at the other ends of the electrodes 110 and 112 close to the output facet 106 of the EO material 102. If the phase velocities of the optical and modulating waves are equal to each other, a portion of an optical wavefront of the optical wave experiences the same instantaneous modulating electrical field. This corresponds to the field which that portion of the optical wavefront encounters at the input facet 104. Such a match in the phase velocities is desirable since the efficiency of modulation is maximum as described hereinafter.

However, material dispersion often adversely affects or prevents this desired phase match condition. Phase velocities of electrical and optical waves are different in most cases. The electrical modulation wave travels at a phase

2

velocity v_{mod} usually lower than the phase velocity v_{opt} of the optical wave. In an integrated modulator, this mismatch in phase velocities is primarily caused by the dispersion of the electro-optic material of the substrate. The modulation wave and the optical wave then experience different indices of refraction due to their different frequencies. This can limit the EO modulator to operate at higher modulation frequencies and significantly reduce the modulation depth.

The phase modulation depth δ of the traveling-wave modulator 100 can be approximately expressed as

$$\delta = \beta L \frac{\sin(\omega_{mod} \Delta L)}{\omega_{mod} \Delta L}, \quad (1)$$

where β is a factor linearly proportional to the magnitude of the electrical field of the modulation wave and the electro-optic coefficient of the crystal and Δ is a phase mismatch parameter:

$$\Delta = \frac{1}{2} \left(\frac{1}{v_{mod}} - \frac{1}{v_{opt}} \right). \quad (2)$$

The phase modulation depth δ increases proportionally with the interaction length L. In addition, the phase modulation depth δ is dependent on a phase mismatch parameter Δ and reaches at or near a maximum value when $\Delta=0$, i.e., the phase velocities v_{mod} and v_{opt} are equal to each other. Otherwise, the phase modulation is reduced by a factor due to the phase mismatch between the optical wave and the electrical modulation wave. An ideal phase match requires $v_{mod}=v_{opt}$. This is usually difficult to achieve in practical devices. In practice, the phase match condition may be considered satisfactory if

$$\omega_{mod} \Delta L \ll 1. \quad (3)$$

Another parameter which sets the performance of the EO modulators is the sensitivity of modulation. It is desirable to achieve a maximal phase modulation depth δ with a smallest possible drive signal power for a given electro-optic material. This can be accomplished, for example, by increasing the interaction length L of the electro-optic material while maintaining the phase match condition in Equation (3). In addition, choosing an electro-optic material with a large electro-optic coefficient can reduce the drive power required for the modulator.

The phase match condition of the optical wave and the electrical modulation wave can be achieved by implementing a traveling-wave configuration in an EO modulator. In general, this can be done by either reducing the phase velocity of the optical wave or speeding up the electrical wave. One way to accomplish the former is disposing bends with a higher refractive index in the waveguide. However, loss by light scattering in the bends is often unacceptably high, thus making this technique impractical in many applications.

There have been two different approaches in increasing the effective phase velocity of the electrical wave to match that of the optical wave.

One approach utilizes a buffering layer of a low index insulator, such as a SiO_2 layer, between the electro-optic crystal and an electrode wherein the electrical wave propagates. See, Gopalakrishna et al., "40 GHZ Low Half Wave Voltage Ti:LiNbO₃ Intensity Modulators", Electronic Letters, Vol. 28, pp. 2056-2068 (1995). This buffering layer lifts the fields out of the crystal and effectively reduces the index of refraction experienced by the electrical wave. As a result, the phase velocity of the electrical wave is increased. This velocity matching technique can be effective. However,

field penetration to the crystal is reduced and consequently the modulation sensitivity is compromised. Another limitation of this approach is that a large and uniform buffering layer is usually difficult to grow and can be expensive.

A second alternative approach is to increase the speed of the electrical wave on the average rather than uniformly in the electrical path. Therefore, the phase velocity of an electrical wave may not have been equal to the optical phase velocity. However, on the average, the phase mismatch between the electrical wave and the optical wave is minimized. One implementation of this approach uses a plurality of small and separate electrodes to form an electrode array along each side of the optical path (e.g., the optical waveguide in an integrated system) instead of having a single-piece electrode on each side of the optical path in the electro-optic material. The modulator splits the electrical modulation wave accordingly into a plurality of portions with each being fed to one of the small electrodes. Multiple electrical paths from the modulation signal source to each small electrode are formed in a way such that the initial phase of each portion of the input electrical modulation wave at a corresponding small electrode matches the phase of the optical wave. A small phase mismatch develops as each portion of the electrical wave propagates along each small electrode. However, the average phase mismatch for the entire electrode array is substantially eliminated. This is because, at least in part, the length of each small electrode is small compared to the interaction length.

U. S. Pat. No. 5,076,655 to Bridges describes such a system which uses an antenna array to implement the electrode array. An electrical modulation wave is used to illuminate the antenna array at a selected angle with respect to the array so that a modulation electrical wave arriving at each antenna has a different phase delay and matches the phase of the optical wave at that antenna. U.S. Pat. No. 5,291,565 to Schaffner et al. discloses another system to implement the above technique.

This second approach for speeding up the electrical wave may also be limited in several respects. For example, splitting the electrical modulation wave into multiple waves (e.g., N waves) reduces the voltage on each electrode in the array by a factor of $N^{1/2}$, thereby decreasing the modulation sensitivity. In addition, the structure of the RF electrode array is complex and accordingly the construction of the entire modulator becomes complicated.

SUMMARY OF THE INVENTION

The present disclosure describes a reflection-mode electro-optic modulator using a traveling electrical modulation wave. According to one embodiment of the invention, a reflective electro-optic modulator includes an electro-optic material having an optical path therein to transmit an optical signal, an optical reflector disposed relative to the electro-optic material to reflect the optical signal to retrace the optical path, and at least one electrode elongated in a direction parallel to the optical path and having a first end for receiving said electrical modulation signal and a second end with a signal reflector for reflecting said electrical modulation signal to retrace the electrode. The electrode has an electrode length shorter than the optical length by an appropriate amount to effect a phase matching between the reflected electrical modulation signal and the reflected optical signal.

The location at which the electrical and optical signals have the same phase delay may be preferably configured to be near or at the second end of the electrode to minimize the length of the electro-optic material. One advantage of this

configuration is that the interaction length is nearly doubled for a given electro-optic crystal compared to the interaction length of a conventional traveling-wave EO modulator. This increased interaction length can significantly improve the modulation sensitivity.

Another advantage includes maintaining the simplicity of the traveling-wave geometry in achieving rephasing of the optical and electrical signals.

Yet another advantage is rephasing without splitting the input electrical modulation power. Comparing with conventional two-segment electrode modulators, this automatically results in a gain of 3 dB.

The reflective rephasing scheme can be implemented in a variety of EO modulators, including but not limited to, Mach-Zehnder modulators and directional couplers. Using reflective rephasing in linearized modulators, e.g., a dual series Mach-Zehnder modulator, may significantly improve the dynamic range.

These and other aspects and advantages of the invention will become more apparent in light of the following detailed description, including the accompanying drawings and appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a block diagram of a typical conventional traveling-wave electro-optic light modulator.

FIG. 2A is diagram showing one embodiment of a phase modulator in reflective traveling wave configuration.

FIGS. 2B and 2C are cross sectional diagrams showing two EO modulators using an ion-milled trench to increase the effective phase velocity of a traveling electrical wave.

FIG. 3A is a diagram showing an embodiment of a Mach-Zehnder modulator in reflective traveling wave configuration.

FIG. 3B is a diagram showing an embodiment of a linearized dual series Mach-Zehnder modulator in reflective traveling wave configuration.

FIG. 3C is a chart showing frequency dependence of the dynamic range of a conventional linearized Mach-Zehnder modulator and a reflective linearized Mach-Zehnder modulator.

FIG. 4A is diagram of a directional coupler in reflective traveling wave configuration.

FIG. 4B is diagram of a linearized directional coupler in reflective traveling wave configuration.

DETAILED DESCRIPTION OF THE INVENTION

1. Traveling-Wave Reflective Phase Modulator

FIG. 2A shows one embodiment 200 of the traveling-wave reflective phase modulator in accordance with the invention. This is an integrated modulator formed on an electro-optic material 202 (e.g., LiNbO₃ and GaAs crystals) having two opposing facets 206 and 208 that are preferably parallel to each other. An optical waveguide 210 is formed in the EO material 202 to provide an electrically-controlled optical conduit, preferably along a direction substantially perpendicular to facets 206 and 208. The facet 206 is transparent and used for transmitting an optical wave. Facet 208 is reflective so that an input optical wave to the waveguide 210 is reflected back toward the input, thereby traversing the waveguide 210 twice. This can be achieved by applying a reflective coating on the inner surface of the facet 208. Alternatively, an optically reflective surface such as a mirror may be used to reflect the optical wave.

Two electrodes 220 and 222 are formed adjacent to the optical waveguide 210. The electrodes 220 and 222 have respective first ends 221, 223 on the side of the facet 206, and second ends 225, 227 on the other end. Preferably, the electrodes 220 and 222 are substantially parallel to each other and are shorter than the length of the EO material 202 between facets 206 and 208. The electrodes 220 and 222 can be used to provide a modulation electrical field at least in a portion of the optical waveguide 210 extending from the first end near the facet 206 to the second end near the facet 208. Flat stripline conductors, for example, may be formed on the surface of the electro-optic material 202 as the electrodes 220 and 222.

A signal generator 230 is electrically connected to electrodes 220 and 222 at the first ends 221, 223 near the facet 206 and initiates a traveling modulation wave. Each electrode has a signal reflector 240 at the second end (225 or 227) close to the facet 208 in order to reflect the modulation wave and form a reflected modulation wave with the same magnitude and phase as the original incident modulation wave but propagating in the opposite direction. An open circuit transmission line termination, for example, may be used as the signal reflector 240.

In this embodiment, an optical wave launched in the device travels along an optical path from the facet 206, to the reflective facet 208, and back to the facet 206. This optical path is longer than an electrical path of the traveling modulation wave which is from the ends 221, 223, to the other ends 225, 227, and back to the ends 221, 223 of the electrodes 220 and 222. Hence, the reflected optical wave is delayed in phase relative to the reflected electrical wave. This phase delay can be used to phase match reflected optical and electrical waves at any location between the facet 206 and the terminators 240 covered by the electrodes. The optical phase delay may be adjusted by selecting the length of the EO material 202 so that a reflected optical wave is in phase with a reflected electrical wave at a selected location.

One phase-matching location for the reflections is at the second ends 225, 227 of the electrodes 220 and 222 close to the reflecting facet 208. Assume the length of the electrodes is L_1 and the EO material 202 is longer than the electrodes by an amount L_2 . The total length of the crystal is thus $(L_1 + L_2)$. Phase matching at the second ends 225, 227 of the electrodes 220 and 222 of optical and electrical waves that are simultaneously launched at the facet 206 requires:

$$\frac{L_1}{c/n_{\text{mic}}} = \frac{2L_2 + L_1}{c/n_{\text{opt}}}, \quad (4)$$

where c is the speed of light in vacuum, n_{mic} is the refractive index at a frequency of the electrical wave, n_{opt} is the refractive index at a frequency of the optical wave, respectively. The desired extra length L_2 of the EO material 202 can thus be determined:

$$L_2 = \frac{1}{2} \left(\frac{n_{\text{mic}}}{n_{\text{opt}}} - 1 \right) L_1. \quad (5)$$

For LiNbO_3 , $n_{\text{opt}}=2.2$, $n_{\text{mic}}=4.0$, thus, $L_2=0.41 L_1$. Therefore, a LiNbO_3 crystal should be about 41% longer than the rf electrodes. For simplicity of discussion, it is assumed that the first ends 221, 223 of the electrodes 220 and 222 are so close to the facet 206 and the phase delay associated with the spacing between the facet 206 and the first ends 221, 223 can be neglected.

Equation (5) indicates a minimum extra length of an EO crystal for phase matching in such a reflective phase modulator 200. Phase matching at other locations requires an

appropriate amount of increase over the minimum crystal length. For example, if the phase matching location is x away from the reflective terminator 240 in the electrodes, the extra length L_2 is then given by:

$$L_2 = \frac{1}{2} \left(\frac{n_{\text{mic}}}{n_{\text{opt}}} - 1 \right) L_1 + \frac{1}{2} \frac{n_{\text{mic}}}{n_{\text{opt}}} x \quad (6)$$

The extra real estate, L_2 , of the EO material 202 may be utilized for placing one or more bias electrodes that are often necessary for proper operation of a modulator, provided that there is no capacitive coupling onto the bias electrodes.

One advantage of the system 200 is the significantly improved utilization of the EO material which results from making the effective interaction length double that of the electrode length. This improves the compactness of the device and reduces the manufacturing cost since EO crystals are usually expensive.

One may readily appreciate other advantages of the embodiment 200 of the invention. For example, the embodiment 200 has a simple construction. The simple traveling-wave geometry is maintained without additional complexity in construction geometry and electrical wiring. Such a configuration can facilitate the device integration, improve reliability and further reduces the cost.

For another example, the embodiment 200 accomplishes the rephasing by effectively using a single-piece electrode on each side of the optical waveguide 210 in the EO material 202 to function as two identical electrodes connected in series. Thus, the embodiment 200 can be viewed as a special case of multiple segment electrodes. However, unlike convention systems with multiple segment electrodes, the embodiment 200 does not partition the rf power from the signal generator 230 into multiple parts. Therefore, a 3-dB gain is automatically achieved in the embodiment 200 over a conventional system having two electrode segments.

The embodiment 200 may be combined with other velocity matching techniques to achieve light modulation at an increased frequency. Velocity matching techniques reduce the inherent velocity dispersion in the crystal. For example, a large SiO_2 buffering layer may be disposed between the electrodes and the crystal to reduce the microwave index of refraction and thus the velocity dispersion.

Another technique is the ion milling of trenches in the LiNbO_3 . The milling removes a portion of the high index EO material outside the optical waveguide or optical path but adjacent to the electrodes. This portion of EO material would otherwise be exposed to the electric field and further reduce the velocity of the electrical wave. FIG. 2B shows a directional coupler formed in an EO substrate 250 with an ion-milled structure. A portion of the EO material is removed by ion milling to form a trench 252 between optical waveguides 257 and 258. Thus, an electrical field between the waveguides produced by two electrodes 254 and 258 will interact with air rather than the EO material with higher refractive index. This reduces the effective refractive index of the traveling electric wave and thereby speeds up the phase velocity. Due to the proximity of the two waveguides 257 and 258, the electrical field produced by the electrodes 254 and 258 (indicated by the dashed arrow lines) is primarily in the vertical direction within the two waveguides 257 and 258. The EO crystal may be cut to be sensitive to the vertical field component (e.g., z-cut). FIG. 2C shows another example where a waveguide 262 formed in an EO substrate 260. Two trenches 264 and 266 are formed on both sides of the waveguide 262 to reduce the effective refractive index by the traveling electrical wave produced by the two electrodes 268a and 268b. Since the electrical field within

the waveguide 262 is primarily horizontal, the EO crystal may also be cut accordingly to increase the electro-optic effect (e.g., x-cut). This configuration may be used as a phase modulator or one modulation branch in a Mach-Zehnder modulator.

Partial velocity matching may be used on a traveling wave reflective modulator in accordance with the present invention to increase operating bandwidth. A difference in the refractive indices experienced by the electrical and optical waves can be reduced by various techniques but it is usually difficult to null this difference completely. For example, a modulator design might meet the specification for sensitivity but might also need a 3-dB point that occurs at twice the frequency. Here, one could attempt to reduce the refractive index difference by a factor of two through a velocity matching technique.

It should be understood that the embodiment 200 of FIG. 2 is only one example, and more specifically an integrated example, of the traveling-wave reflective phase modulator. Many other configurations are possible. For example, a lumped reflective modulator may be constructed by using either a mirror or a reflective coating to reflect the optical wave back to the EO crystal.

2. Traveling-Wave Reflective Mach-Zehnder Modulators

FIG. 3A shows an integrated Mach-Zehnder modulator 300 using the reflection configuration in accordance with the present invention. The modulator 300 is formed on a piece of electro-optic material 302 (e.g., LiNbO₃ crystal) having two opposing facets, the input facet 306 and reflective facet 308 that are preferably parallel to each other. The reflective facet 308 may be formed by having reflective coating thereon.

Three optical waveguides 310, 312, and 314 are formed in the EO material 302. The optical waveguide 310 may be substantially perpendicular to the input facet 306 and serve as the optical input and output of the modulator 300. The waveguides 312 and 314 are substantially parallel to each other and have substantially the same optical path length. A waveguide coupler 311 couples the waveguide 310 to the first ends of waveguides 312 and 314 so that an input optical wave from the waveguide 310 can be split into two substantially identical portions to waveguides 312 and 314. Both waveguides 312 and 314 are preferably perpendicular to the reflective facet 308 and terminated at the facet 308 so that an optical wave traveling towards the facet 308 in the waveguide will be reflected to propagate away from the facet 308.

Three substantially identical and elongated electrodes 320, 322, and 324 are disposed in the EO substrate 302 parallel to the optical waveguides 312 and 314. A modulation signal generator 330 is electrically connected to the first ends of the electrodes close to the coupler 311 and initiates a traveling modulation electrical wave into the electrodes. Each electrode has a signal reflector 240 (e.g., an open circuit transmission line terminator) at a second end close to the reflective facet 208.

The EO substrate 302 is extended by a predetermined length L₂ beyond the signal reflector 240 of the electrode in order to match the phases of reflected electrical and optical waves at a selected location as described previously. The extra substrate area may be used to place bias electrodes 340.

The Mach-Zehnder modulator 300 is in effect a "folded" conventional Mach-Zehnder modulator. An input optical wave is fed through the optical waveguide 310 and split into two different signals in waveguides 312 and 314. The light modulation is implemented by launching a traveling elec-

trical signal into the electrodes 320, 322, and 324. The split optical signals are reflected back from the reflective facet 308 and recombined at the coupler 311. The resultant optical signal is then sent out by the waveguide 310.

The reflective scheme in accordance with the invention may be used to construct linearized electrooptic modulators. Linearization in EO modulators reduces harmonic signals and intermodulation products in order to improve the modulation dynamic range. FIG. 3B shows an embodiment of a linearized dual series Mach-Zehnder modulator (DSMZ) 350. The DSMZ 350 has a complete conventional Mach-Zehnder optical path having a first optical coupler 311, two waveguides 312 and 314, and a second optical coupler 313 which is terminated by a reflective facet of the OE substrate. The reflective scheme allows the single Mach-Zehnder optical path to function effectively as two Mach-Zehnder modulators: a first modulator using the first coupler 311 as its input and the second coupler 313 as its output, a second modulator using the second coupler 313 as its input and the first coupler 311 as its output. The two modulators are connected in series: the input of the second modulator is the reflection of the output of the first modulator. Note that both modulators share the same set of bias electrodes 360 with the same electrical voltage. The linearization of a DSMZ with identical bias voltages on the two modulators has been demonstrated. See, Bets and O'Donnell, "Microwave analog optical links using suboctave linearized modulators", in IEEE Photonics Technology Letters, Vol. 8(9), pp. 1273-1275 (1996), which is incorporated herein by reference.

The rephased reflected optical wave is especially useful in linearized electrooptic modulators. This is in part because the linearization schemes use sensitive cancellation techniques and a small amount of velocity mismatch may cause severe degradation in performance. For example, a Linearization scheme may yield an extra 20 dB of dynamic range at low frequencies, but the improvement can be quickly lost as the frequency increases due to phase mismatch. The effects are much more pronounced than the reduction in gain, which is normally considered the limiting factor for operation at high frequencies (e.g., larger than 1 GHz).

The advantage of the invention in this aspect may be illustrated by FIG. 3C, which shows the calculated dynamic range for a LiNbO₃ DSMZ modulator with a set of typical design parameters. The line 370 at 110 dB is a reference line of a non-linearized modulator. It does not degrade with frequency because while the gain decreases with frequency, the distortion terms attenuate at the same rate. Curve 372 represents the dynamic range as a function of frequency for a linearized modulator with a single-segmented traveling wave electrode. An initial improvement in the dynamic range of 24 dB at low frequencies falls off dramatically with increasing frequency. Similar analysis on other conventional linearized modulators shows a similar rapid roll-off in dynamic range when there is no rephasing of the electrical and optical waves.

The reflective linearized DSMZ modulator 350 of FIG. 3B, however, shows a much improved dynamic range at about 133 dB that remains substantially unchanged with the frequency (curve 374). Since the reduction in dynamic range is a function of the phase error at the second Mach-Zehnder modulator, the rephasing between the electrical and optical waves at the input of the second Mach-Zehnder minimizes the roll-off of the enhanced dynamic range. The calculation shows that the gain of the reflective linearized modulator 350 is limited. The 3-dB gain point in the calculation is about 8 GHZ. The fact that one rephasing is sufficient to

make the reflective linearized modulator robust to velocity mismatch demonstrates its commercial potential for linearized electrooptic modulators.

3. Traveling-Wave Reflective Directional Coupler Modulators

Directional couplers are well known in the art. The inventor of the present invention recognized that a reflective directional coupler can be constructed using the reflective scheme of the invention to achieve a more compact configuration and improved coupling efficiency over the conventional couplers.

FIG. 4A shows one embodiment 410 of such a directional coupler. An EO substrate has two waveguides 412 and 414 formed at a close proximity in a substantially parallel fashion with respect to each other to facilitate energy coupling therebetween. The waveguides 412 and 414 have input/output waveguide terminals 413 and 415 that are apart from each other to minimize the energy coupling. The waveguide 412 may be used as the input of the coupler 410. The performance of the directional coupler 410 may be characterized by a transfer function, H:

$$H \approx \sin\left[\left(kL\right)\sqrt{1 + \left(\sqrt{3} - \frac{V}{V_r}\right)}\right]. \quad (7)$$

where k is the coupling coefficient, L is the length of the modulator, V is the voltage amplitude and V_r is normalization constant in volts. The characteristic electrical length, (kL), of the waveguides can be chosen by design to obtain an electrical path length equivalent to a phase of either $(kL)=\pi/2$ or π . If $(kL)=\pi/2$, all of the input optical power is coupled to the second guide 414 and comes out at the terminal 415. If $(kL)=\pi$, then all of the input power will come out of the input guide 412 at the terminal 413. The choice of the characteristic electrical length is determined by a specific requirement of a device or system.

For example, in a remote application, it may be desirable to have only one fiber and to locate a laser away from the modulator. Therefore, the characteristic electrical length of the electrical path should be set to π . However, other applications may require $\pi/2$ and eliminate the need for an optical circulator. Note also that the bias electrodes 420 conveniently fit in the space allocated for delaying the optical wave.

FIG. 4B shows a linearized reflective directional coupler using the reflective scheme of the invention. The linearization mechanism is used here to achieve an improved operation bandwidth over a non-linearized directional coupler of FIG. 4A. It is known that each channel in a linearized directional coupler needs two bias electrodes each with an electrical length of $\pi/4$ and a rf electrode of a length of $\pi/2$ for applying a modulation electrical signal. With standard manufacturing dimensions in LiNbO_3 , this requirement leads to a long modulator using a conventional linearized directional coupler configuration. In practice, fabrication of such modulators with long EO crystals may be difficult and expensive. The reflective modulator of FIG. 4B, however, can have the same electrical lengths while being only half as long.

Although the present invention has been described in detail with reference to a few embodiments, one ordinarily skilled in the art to which this invention pertains will appreciate that various modifications and enhancements may be made without departing from the scope and spirit of the invention. For example, different geometries other than those disclosed may be used. The waveguide(s) and the electrodes for carrying the modulation signals may be bent

or curved. The electrodes may be formed by multiple separated conductors disposed along the waveguide(s) extending from the first end to the second end in which the effective length of the electrode(s) may be electrically changed to adjust the phase matching.

In addition, a variety of electro-optical materials may be used other than those mentioned, including but not limited to, Gallium Phosphide (GaP), Lithium Tantalate (LiTaO_3), KDP, BaTiO_3 , and Potassium Niobate (KNbO_3). Some commonly used EO materials that can be used in the present invention are listed in Table 9.2 in Chapter 9 of Optical Electronics, 4th edition, by A. Yariv, Saunders College Publishing (1991), which is incorporated herein by reference.

These and other variations are intended to be encompassed by the invention which is defined in the following claims.

What is claimed is:

1. An electro-optic modulator, comprising:
an electro-optic material having an index of refraction responsive to an electrical modulation signal and defining an optical path of a selected optical length to transmit an optical signal that traces said optical path;
an optical reflector, disposed relative to said electro-optic material to reflect said optical signal to retrace said optical path;
at least one electrode disposed relative to said electro-optic material and configured to provide an electrical conduit for said electrical modulation signal along said optical path, said electrode having a first portion to receive said electrical modulation signal; and
a signal reflector disposed at a second portion of said electrode spaced from said first portion and configured to reflect said electrical modulation signal back to said first portion, wherein said electro-optic material and said electrode are configured to effect a phase matching between said electrical modulation signal reflected by said signal reflector and said optical signal reflected by said optical reflector.
2. An electro-optic modulator as in claim 1, wherein said optical reflector is a reflective coating formed on a facet of said electro-optic material.
3. An electro-optic modulator as in claim 1, wherein said optical reflector is a mirror located relative to said electro-optic material.
4. An electro-optic modulator as in claim 1, wherein said signal reflector is an open circuit transmission terminator.
5. An electro-optic modulator as in claim 1, wherein said electrode is elongated in a direction substantially parallel to said optical path and has an electrode length shorter than said optical length by a specified length amount.
6. An electro-optic modulator as in claim 5, wherein said length amount is so chosen that said electrical modulation signal reflected by said signal reflector and said optical signal reflected by said optical reflector have the same phase at said second portion of said electrode.
7. An electro-optic modulator as in claim 1, wherein said optical path is defined by an optical waveguide formed within said electro-optic material.
8. An electro-optic modulator as in claim 1, wherein said electrode is formed of a single piece of conductive material on said electro-optic material.
9. An electro-optic modulator as in claim 1, wherein said first and second portions in said electrode are respectively two ends of said electrode.
10. An electro-optic modulator as in claim 1, wherein said electrode includes a plurality of separate conductive elements.

11

11. An electro-optic modulator, comprising:
 an electro-optic crystal of a prespecified crystal length
 having a first transmissive facet receiving and exporting
 an optical signal, and a second reflective facet
 substantially parallel to said first transmissive facet,
 said electro-optic crystal configured to define an optical
 waveguide to form an optical conduit between said first
 and second facets; and
 a pair of substantially parallel electrode stripes formed on
 said crystal and located on each side of said optical
 waveguide, each of said electrode stripes having a first
 portion close to said first transmissive facet which
 receives an electrical modulation signal and a second
 portion close to said second reflective facet which has
 an open circuit transmission terminator to reflect said
 electrical modulation signal towards said first portion,
 a distance between said first and second portions defining
 an electrode length.
 wherein said crystal length is longer than said electrode
 length by an amount such that said optical signal
 reflected from said second facet undergoes the same
 phase delay at a phase-match location between and
 including said first and second portions as said electrical
 modulation signal that is reflected from said trans-
 mission terminator.

12. An electro-optic modulator as in claim 11, wherein
 said phase-match location is at said second portion.

13. An electro-optic modulator as in claim 11, wherein
 said first and second portions are respectively located at two
 ends of said electrode stripes.

14. An electro-optic modulator as in claim 11, wherein
 said crystal is selected from a group consisting of LiNbO₃,
 GaAs, GaP, LiTaO₃, KDP, BaTiO₃, and KNbO₃.

15. An electro-optic modulator as in claim 11, further
 comprising a plurality of bias electrodes located on said
 crystal between said second reflective facet and said second
 portion of said electrode stripes and configured to provide
 bias electrical signals to said waveguide.

16. An electro-optic modulator as in claim 11, wherein
 said crystal is configured to have a second optical waveguide
 disposed with respect to said optical waveguide in parallel
 between said pair of electrode stripes to form a second
 optical conduit between said first and second facets in said
 crystal, wherein said optical waveguides are in a close
 proximity to each other to effect an energy coupling there-
 between to form a directional coupler modulator.

17. An electro-optic modulator as in claim 16, further
 comprising bias electrodes located on said crystal near said
 optical waveguides between said second portion of said
 electrode stripes and said second facet of said crystal, said
 bias electrodes operating to effect a linearization mechanism
 in said directional coupler modulator.

18. An electro-optic modulator that effects a Mach-
 Zehnder modulator, comprising:

an electro-optic crystal of a prespecified crystal length
 having a first transmissive facet receiving and exporting
 an optical signal, and a second reflective facet
 substantially parallel to said first transmissive facet;
 first and second optical waveguides formed in said crystal
 and configured to have a first end and a second end, said
 second end terminated at said second reflective facet;
 an optical coupler formed in said crystal and configured to
 combine said first ends of first and second waveguides
 to form a common waveguide terminated at said first
 transmissive facet, said common waveguide receiving
 and splitting said input optical signal into first and

12

second optical signals in said first and second
 waveguides, respectively;
 three substantially parallel electrode stripes formed on
 said crystal and disposed relative to said first and
 second waveguides so that one of said electrode stripes
 is between said first and second waveguides and the
 other two electrode stripes are respectively located on
 the other side of said first and second electrode stripes,
 each of said electrode stripes having a first portion
 close to said first transmissive facet which receives an
 electrical modulation signal and a second portion close
 to said second reflective facet; and
 an open circuit transmission terminator connected to said
 second portion of each of said electrode stripes to
 reflect said electrical modulation signal towards said
 first portion,
 wherein said crystal and said electrode stripes are con-
 figured to cause said first and second optical signals
 reflected from said second reflective facet to have a
 phase delay at a phase-match location between and
 including said first and second portions equal to said
 electrical modulation signal that is reflected from said
 transmission terminator.

19. An electro-optic modulator as in claim 18, wherein
 said phase-match location is at said second portion.

20. An electro-optic modulator as in claim 18, wherein
 said first and second portions are respectively located at two
 ends of said electrode stripes.

21. An electro-optic modulator as in claim 18, wherein
 said crystal is selected from a group consisting of LiNbO₃,
 GaAs, GaP, LiTaO₃, KDP, BaTiO₃, and KNbO₃.

22. An electro-optic modulator as in claim 18, further
 comprising a plurality of bias electrodes located on said
 crystal between said second reflective facet and said second
 portion of said electrode stripes and configured to provide
 bias electrical signals to said first and second waveguides.

23. An electro-optic modulator that effects two Mach-
 Zehnder modulators optically connected in series, comprising:
 an electro-optic crystal of a prespecified crystal length
 having a first transmissive facet receiving and exporting
 an optical signal, and a second reflective facet;
 first and second optical waveguides formed in said crystal
 and configured to have a first end and a second end;
 a first optical coupler formed in said crystal and config-
 ured to combine said first ends of first and second
 waveguides to form a first common waveguide termin-
 ated at said first transmissive facet, said first common
 waveguide receiving and splitting said optical signal in
 said first and second waveguides, respectively;

a second optical coupler formed in said crystal and config-
 ured to combine said second ends of first and second
 waveguides to form a second common waveguide termin-
 ated at said second reflective facet, said second common
 waveguide receiving and splitting a reflected optical signal from said second reflective
 facet in said first and second waveguides, respectively;

three substantially parallel electrode stripes formed on
 said crystal and disposed relative to said first and
 second waveguides so that one of said electrode stripes
 is between said first and second waveguides and the
 other two electrode stripes are respectively located on
 the other side of said first and second electrode stripes,
 each of said electrode stripes having a first portion
 close to said first transmissive facet which receives an